DOCUMENT RESUME

ED 118 359 SE 019 681

AUTHOR	Hadar, Nitsa Boneh
TITLE	Children's Conditional Reasoning: An Investigation of Fifth Graders' Ability to Learn to Distinguish Between Valid and Fallacious Inferences.
PUB DATE	75
NOTE	420p.; Ph.D. Dissertation, University of California, Berkeley
EDRS PRICE	MF-$0.83 HC-$22.09 Plus Postage
DESCRIPTORS	Curriculum; Doctoral Theses; *Educational Research; *Elementary Education; *Elementary School Science; *Learning Processes; Logic; *Logical Thinking; Science Education; Thought Processes
IDENTIFIERS	*Conditional Reasoning; Research Reports

ABSTRACT

This study was conducted to determine if fifth-grade students can significantly improve their use of logical analysis through a suitable instructional unit taught under ordinary classroom conditions. Concrete teaching materials were developed to familiarize students with the distinction between the valid inference patters--Modus Ponendo Ponens and Modus Tollendo Tollens (MP, MT), and the fallacious ones--Affirming the Consequent and Denying the Antecedent (AC, DA). No formal rules were taught. The experimental unit was implemented 4 to 5 times a week for 23-25 sessions, by 4 fifth-grade teachers in their ordinary classes. The teachers participated in a 12-hour pretraining workshop. A pretest/posttest, treatment/no treatment design was applied to assess resulting improvement in students' conditional reasoning ability. The sample consisted of 210 fifth-grade students in a suburban area, 104 in 4 experimental classes and 106 in 4 control classes. Experimental and control group pretest performance levels did not differ, but there was a significant difference in the posttest means. There was no significant change in the control group's pretest and posttest performance levels on any logical form, or for the experimental group's on MP and MT. However, on AC and DA the two groups' gain scores were significantly different. (MLH)

****
* Documents acquired by ERIC include many informal unpublished *
* materials not available from other sources. ERIC makes every effort *
* to obtain the best copy available. Nevertheless, items of marginal *
* reproducibility are often encountered and this affects the quality *
* of the microfiche and hardcopy reproductions ERIC makes available *
* via the ERIC Document Reproduction Service (EDRS). EDRS is not *
* responsible for the quality of the original document. Reproductions *
* supplied by EDRS are the best that can be made from the original. *
****

U.S. DEPARTMENT OF HEALTH
EDUCATION & WELFARE
NATIONAL INSTITUTE OF
EDUCATION

THIS DOCUMENT HAS BEEN REPRODUCED EXACTLY AS RECEIVED FROM THE PERSON OR ORGANIZATION ORIGINATING IT. POINTS OF VIEW OR OPINIONS STATED DO NOT NECESSARILY REPRESENT OFFICIAL NATIONAL INSTITUTE OF EDUCATION POSITION OR POLICY

SE

ED118359

CHILDREN'S CONDITIONAL REASONING
AN INVESTIGATION OF FIFTH GRADERS' ABILITY TO LEARN TO DISTINGUISH BETWEEN VALID AND FALLACIOUS INFERENCES

NITSA BONEH HADAR

The Regents of the
University of California

GRADUATE DIVISION
UNIVERSITY OF CALIFORNIA, BERKELEY

1975

SE 019 681

Children's Conditional Reasoning: An Investigation of Fifth Graders' Ability to Learn to Distinguish between Valid and Fallacious Inferences

By

Nitsa Boneh Hadar

B.S. (Hebrew University, Jerusalem, Israel) 1964
M.S. (Technion, Israel Institute of Technology, Haifa) 1970

DISSERTATION

Submitted in partial satisfaction of the requirements for the degree of

DOCTOR OF PHILOSOPHY

in

Science/Mathematics Education

in the

GRADUATE DIVISION

of the

UNIVERSITY OF CALIFORNIA, BERKELEY

Approved: Leon Henkin
Lawrence F. Lowery
Robert Karplus

Committee in Charge

....

ACKNOWLEDGMENTS

My debt of gratitude is extensive, for the assistance of many people who contributed to the completion of this study.

I am thankful to Leon Henkin, Professor of Mathematics for devoting endless time to critically consider each and every step of my work. The constant availability of his assistance was the greatest source of encouragement I could get.

My sincere appreciation goes to Robert Karplus, Professor of Physics, for his very close supervision and unlimited feedback.

In addition I would like to acknowledge the contributions of the following individuals:

Professor Lawrence Lowery, who was always supportive and provided valuable editorial comments.

Professor Leonard Marascuilo, who advised me in statistical analysis.

Dr. R. Rosenquist, superintendent of Albany Unified School District; E. Allison, Curriculum Specialist of Moraga School District in California; the principals and fourth and fifth grade teachers in these districts who facilitated the experimental work by organizing and carrying it out.

D. McEntyre, colleague in the Graduate Group for Science and Mathematics Education, for volunteering her time to prepare computer programs.

K. Fairwell, of the Lawrence Hall of Science who invested much effort in editing the manuscript; and J. Henke, who very efficiently typed it.

Finally, I am deeply indebted to my children, who grew up independent by necessity, and to my husband, who sacrificed three years of his own career for me. To him I dedicate this dissertation.

This study was partially supported by an International Fellowship from the American Association of University Women and by NSF grants no. GW 7659, PES 74-018950.

CHILDREN'S CONDITIONAL REASONING:
An Investigation of Fifth Graders' Ability to Learn to Distinguish between Valid and Fallacious Inferences

ABSTRACT

Nitsa Hadar

This study stemmed from a desire to redress the distorted view of mathematics in the elementary curriculum, created by the current imbalanced emphasis on computational rules and some applications, but very little logical analysis. The study intends to show that fifth-grade students can significantly improve their use of logical analysis through a suitable instructional unit taught under ordinary classroom conditions.

Concrete teaching materials were developed, through several trials and revisions, to familiarize students with the distinction between the valid inference patterns -- Modus Ponendo Ponens and Modus Tollendo Tollens (MP, MT), and the fallacious ones -- Affirming the Consequent and Denying the Antecedent (AC, DA). No formal rules were taught.

The experimental unit was implemented four to five times a week for 23-25 sessions, by 4 fifth-grade teachers in their ordinary classes. The teachers participated in a twelve-hour pre-training workshop.

A pretest/posttest treatment/no-treatment design was applied to assess resulting improvement in students' conditional reasoning ability. The sample consisted of 210 fifth graders in a suburban area, 104 in 4 experimental classes and 106 in 4 control classes.

A written group test was developed, through trials and revisions. Test items are formulated with a reasonable hypothetical content. Each item includes two premises: the first a conditional sentence, and the second either its antecedent, its consequent, or the negation of one of these, thus determining the logical form: MP, MT, AC, or DA. The question following the premises is stated positively. MP and MT are answered correctly by "yes" or "no"; AC and DA by "not enough clues" (NEC).

The test contains 32 randomly-ordered three-choice items, eight in each logical form (two of the eight in each of the four possible modes in which negation may or may not occur in the antecedent or consequent). No sentential connective other than negation and conditional appears in the premises. Test/retest reliability was .79.

Experimental and control group pretest performance levels did not differ ($\alpha = .05$). More than 78% of the answers on MP and MT, and fewer than 33.1% on AC and DA, were correct. Overall pretest mean scores were 54.3% and 53.8% for the experimental and control groups respectively.

There was a significant difference ($\alpha = .01$) between the experimental and control groups' posttest overall performance - 74.7% and 55.4%, respectively.

There was no significant change in the control group's pretest and posttest performance levels on any logical form, or for the experimental group's on MP and MT. However, on AC and DA the two groups' gain scores were found significantly different.

Negation mode, unlike logical form, was not found to be independently influential in analyzing test scores, but interacted with logical form.

There was a pretest/posttest increase of 3.5 in experimental group frequency (percentaged) of incorrect NEC answers (MP and MT). As NEC appeared infrequently on the pretest, this increase was interpreted as learning that NEC is an acceptable answer. Separating out this effect from the percentaged frequency of correct NEC answers (AC and DA) left a pretest/posttest average increase of 37.8. This increase was attributed to learning when NEC is correct.

Teachers were excited at the beginning, frustrated in the middle, and felt competent and involved in the project at the end. They felt the teaching should be less condensed. The majority of the students reacted positively to most parts of the experimental unit. However, some thought the unit as a whole was too repetitive and boring.

No correlation was found between learning logic through the experimental unit and standard school achievements as measured by the Stanford Achievement Test (SAT). High, average, and low SAT achievers of the experimental group did not differ significantly in their pretest to posttest gain scores.

Results of the study call for further investigation of the value and usefulness of teaching various parts of logic as an ordinary part of the elementary mathematics curriculum.

Table of Contents

Title — Page

Lists of Diagrams and Tables xii

Chapter 1 THE PROBLEM 1

- Chapter Overview 1
- 1.1 General Goals of the Study 2
  - 1.1.1 The role of logical analysis in general education 2
  - 1.1.2 The role of logical analysis in mathematics and in mathematics education 6
- 1.2 Specific Goals - Conditional Reasoning 9
- 1.3 Mathematical Logic Underlying Conditional Reasoning 11
  - 1.3.1 The language of sentential conditional logic 11
  - 1.3.2 Conditional sentences and their negation modes 13
  - 1.3.3 Conditional sentences - their truth value and English interpretation 13
  - 1.3.4 Sentences equivalent to conditional sentences 15
  - 1.3.5 Relations between conditional sentences 16
  - 1.3.6 Two valid and two nonvalid rules of inference 18
  - 1.3.7 Extending the relevant rules of inference to quantificational logic 21
  - 1.3.8 Algorithmic solutions to validity judgments 23
- 1.4 Psychological Research Underlying Conditional Reasoning 25
  - 1.4.1 The relations between symbolic logic and the psychological reasoning process 25
  - 1.4.2 Readiness 27
  - 1.4.3 Content effects 30

Table of Contents (continued)

Title		Page
	1.4.4 Language effects	31
	1.4.5 Language attainment	36
	1.4.6 Context effects	38
	1.4.7 Negation	39
1.5	Methods Previously Employed in Teaching Logic and Those of the Present Study	42
	1.5.1 Previous attempts to teach logic or conditional reasoning	43
	1.5.2 The present study	50
1.6	Teachers' and Prospective Teachers' Conditional Reasoning	51
1.7	Questions Studied	53
	1.7.1 Objectives	53
	1.7.2 Hypothesis tested	54
	1.7.3 Model for error prediction in conditional reasoning	55
Chapter 2	THE EXPERIMENTAL UNIT	58
	Chapter Overview	58
2.1	The Experimental Unit - An Overview	58
	2.1.1 Principles of the approach	58
	2.1.2 Motivation and final project	64
	2.1.3 The first part	64
	2.1.4 The second part	67
	2.1.5 The third part	71
2.2	Unit Development - Experimenter's Work with Individual Students	72
	2.2.1 Exploratory work	72

Table of Contents (continued)

Title		Page
	2.2.2 Basic lessons learned from the exploratory work	74
2.3	Unit Development - Experimenter's Work with Small Groups	75
	2.3.1 Albany, Spring 1974	75
	2.3.2 Basic lessons learned from this experiment	76
	2.3.3 Berkeley, Summer 1974	77
	2.3.4 Basic lessons learned and further development	78
	2.3.5 Fall 1974, Lawrence Hall of Science	78
2.4	The Pilot Study - Teachers' Conducted Teaching	79
	2.4.1 The sample and how it was secured	79
	2.4.2 Teachers' training	80
	2.4.3 The teaching period	81
	2.4.4 Main lessons learned from the pilot study	82
Chapter 3	THE INSTRUMENT OF MEASUREMENT	84
	Chapter Overview	84
3.1	Development of a Test	84
	3.1.1 Limitations and relations with previously developed tests	84
	3.1.2 Early versions and revisions	90
	3.1.3 From one test to four equivalent tests	95
	a. Equivalence based on negation modes	96
	b. Equivalence based on logical forms	97
	c. Mixed order by type-mate interchange	98
	d. Transformation by contrapositives	99
	e. Transformation of logical types	100

Table of Contents (continued)

Title | Page

3.1.4 Field trial of two equivalent versions 101
3.1.5 Final revision 104
3.2 The Final Form of the Instrument 107
3.2.1 The two versions 107
3.2.2 Validity and reliability 107
3.2.3 Administration 109
3.2.4 Precautions taken in administration of the test 110
Chapter 4 THE PRINCIPAL STUDY 112
Chapter Overview 112
4.1 Research Design 112
4.1.1 Lay-out of the design 112
4.1.2 Why a quasi-experimental design? 114
4.2 The Sample 115
4.3 Course of the Project 117
4.3.1 First contacts 117
4.3.2 The teachers 118
4.3.3 The pre-training workshop 119
4.3.4 The teaching period 120
Chapter 5 RESULTS AND THEIR INTERPRETATION 122
Chapter Overview 122
5.1 Pretest Scores 123
5.1.1 Equivalence of control and experimental groups 123
5.1.2 Internal consistency of the test 126

Table of Contents (continued)

Title | Page

5.2 Posttest Scores 129

5.3 Pretest/Posttest Comparison, a Descriptive Analysis 129

- 5.3.1 Overall comparison 129
- 5.3.2 Logical form subtests results 133
- 5.3.3 The individual's progress 137
- 5.3.4 Negation mode subtests results 144
- 5.3.5 A tempting false conclusion about the effect of negation on conditional reasoning 147

5.4 Statistical Analysis of the Results 152

- 5.4.1 Analysis of variance on pretest scores 152
- 5.4.2 Analysis of variance on gain scores 155

5.5 Did Students Learn Anything Beyond the Legitimacy of Not Enough Clues as an Answer 160

- 5.5.1 Analysis of NEC uses by individual students 160
- 5.5.2 Analysis of average use of NEC per item 163
- 5.5.3 Analysis of total NEC uses 165
- 5.5.4 Analysis of right and wrong Yes/No answers with respect to hypothesized model for error prediction 168

5.6 Standard School Achievement and Sex Differences with Respect to the Results of the Present Study 175

- 5.6.1 Non-correlation with S.A.T. and with sex 175
- 5.6.2 High, average, and low S.A.T. achievers' performance in the present study 177

5.7 Attitudes 179

- 5.7.1 Students' attitudes 179
- 5.7.2 Teachers' reactions 183

Table of Contents (continued)

Title		Page
Chapter 6	SUMMARY AND CONCLUSIONS	187
	Chapter Overview	187
6.1	Summary of Findings	187
	6.1.1 Summary of objectives	187
	6.1.2 Summary of development and design	188
	6.1.3 Summary of results	190
6.2	Relation to Previous Studies and Suggestions for Further Investigation	195
	6.2.1 Test scores and their various interpretations	195
	6.2.2 Negation in the conditional premise	197
	6.2.3 Implementation problems	198
	6.2.4 Levels of regular school achievements and socio-economic class	200
	6.2.5 Content effects	201
	6.2.6 More doubts about the present study	202
Chapter 7	APPENDICES	205
7.1	Final Version of the Experimental Unit - Teachers' Manual and Answered Students' Worksheets	205
7.2	Final Version of the Test	321
7.3a	Test Versions Used in Pilot Study	360
7.3b	Item Profiles for T_1	380
7.4a	Students' Attitude Questionnaire	382
7.4b	Teacher's Overall Evaluation of Experimental Unit in Conditional Reasoning	383
7.4c	Teacher's Evaluation of Activities	389
7.5	A Session by Session Account of the Pretraining Workshop	392
BIBLIOGRAPHY		395

List of Diagrams

		Page
Diagram 1.1	Relations between pairs of conditional sentences	17
Diagram 2.1	A sample of the Pictorial Activity	66
Diagram 2.2	A sample of the Domino Activity	68
Diagram 2.3	A sample chart for a Playing Card Activity	70
Diagram 4.1	Research design	113
Diagram 5.1	Distribution polygon of pretest and posttest total score	132
Diagram 5.2	Control group distribution of pretest and posttest scores on the logical form subtests	135
Diagram 5.3	Experimental group distribution of pretest and posttest scores on the logical form subtests	136
Diagram 5.4	Control group distribution of pretest and posttest scores on the negation modes subtests	145
Diagram 5.5	Experimental group distribution of pretest and posttest scores on the negation modes subtests	146
Diagram 5.6	Experimental group distribution of gain scores for the four logical forms subtests	156
Diagram 5.7	Control group distribution of gain scores for the four logical forms subtests	157
Diagram 5.8	Ranges of yes/no right answers on the decidable part of the test and of yes/no wrong answers on the undecidable part of the test	173

List of Tables

		Page
Table 1.1	Truth Value of Conditional Sentences	14
Table 1.2	Rules of Inference and Their Validity	20
Table 1.3	Rules of Inference Extended	24
Table 1.4	Algorithmic Solution to Relevant Problems of This Study	26
Table 1.5	Expected Errors in Conditional Reasoning	56
Table 3.1	Percentage of Pilot Study Subjects in CTBS Levels	102
Table 3.2	Pilot Study Mean Scores by Test Versions Order	103
Table 3.3	Final Revision of the Test	104
Table 3.4	Test Item Numbers by Type	108
Table 4.1	Basic Data on the Sample	116
Table 4.2	Stanford Achievement Test Scores for Experimental and Control Classes	117
Table 4.3	Basic Data on the Teachers	119
Table 5.1	Experimental vs. Control Groups Pretest and Stanford Achievement Test (S.A.T.) Percentaged Mean Score	124
Table 5.2	Percentages of Experimental and Control Group Students Selecting Each Answer for the Various Pretest Items	125
Table 5.3	Test Item Profiles Based on Pretest Results of 210 Participating Students	127
Table 5.4	Percentages of Experimental and Control Group Students Per Answer for Each Posttest Item	130
Table 5.5	Distribution of Pretest and Posttest Overall Test Scores for the Experimental and Control Groups	131
Table 5.6	Percentages of Students in the Experimental and Control Groups Having Various Scores on Logical Form Subtests in the Pretest and the Posttest	133

List of Tables (continued)

		Page
Table 5.7	Distribution of Pretest versus Posttest Performance Levels on Each Logical Form Subtest	138
Table 5.8	Distribution of Experimental Group Pretest versus Posttest Performance Levels on all Undecidable Items (AC, DA)	143
Table 5.9	Percentages of Students in the Experimental and Control Groups Having Various Scores on Negation Mode Subtests	144
Table 5.10	Total Agreement and Complete Disagreement Among Hypothesized Raters on Rank Order of Hypothetical Ratees	148
Table 5.11	Rank Order of Increasing Difficulty of (i) Logical Forms within Negation Modes and (ii) Negation Modes within Logical Forms	150
Table 5.12	Analysis of Variance on Pretest Scores	152
Table 5.13	Class Means and Standard Deviations for Pretest, Posttest, and Gain Scores	154
Table 5.14	Analysis of Variance on Pretest to Posttest Gain Scores	158
Table 5.15	Distribution of the Difference between Individual Gains on Right Uses of NEC and Wrong Uses of NEC - Experimental Group	161
Table 5.16	Percent of Students Giving the Answer NEC Per Item (Averaged) in Each Logical Form Subtest	163
Table 5.17	Observed Number of Right and Wrong Uses of the Answer "Not Enough Clues" by Experimental Group Students (n=104)	165
Table 5.18	Expected Additional Posttest Wrong Uses of NEC Based upon Various Assumptions on the Portion Due Only to Legitimization of that Answer	168
Table 5.19	Item Profiles for Experimental Group Pretest and Posttest (n=104)	170
Table 5.20	Range of Yes/No Right Answers on MP and MT and Expected Wrong Answers on AC and DA	172
Table 5.21	Average Number of Wrong Answers Per Item in MP and MT	174

List of Tables (continued)

Page

Table 5.22 Correlation Coefficients Between Sex, Stanford Achievement Test Subtests, and Pretest, Posttest and Gain Scores of the Present Study 176

Table 5.23 Pretest, Posttest, and Gain Score Means for Subgroups of the Experimental Group (n=104) by Levels of Achievement on Stanford Achievement Test 178

Table 5.24 Number of Experimental Group Students Who Indicated They Liked Each Activity 179

Table 5.25 Attitude of Experimental Group Students by Volunteering to Go on Learning Logic after the Posttest 181

Table 6.1 Order of Number of Correct Responses for Negation Modes within Logical Forms in O'Brien's (1972) Study and in the Pretest for the Experimental Group (n=104) of the Present Study 198

CHAPTER 1

THE PROBLEM

Chapter Overview

This chapter presents the specific questions that motivated the study. These questions can be approached in two ways: from the top down or from the bottom up.

Starting from the top, the general need for school training in logical analysis, the research problems center on young students learning the logic of conditional reasoning. A discussion of the significance of conditional reasoning to both general and mathematics education appear in sections 1.1 and 1.2.

Starting from the bottom, the earlier studies of logical thinking, the research problems are derived from literature that indicates the need for effective training in this area for both children and teachers. Those parts of mathematical logic underlying inferences from conditional statements, along with a detailed survey of the literature on psychological research in conditional reasoning, are given in sections 1.3 and 1.4.

While these studies illustrate the present state of affairs, they fail to deal with the capability of young children to progress through a proper teaching program. Some experimenters and psychologists have speculated on this question and a few attempts have been made to teach logic in the elementary school. The description of this study's teaching approach is compared to previous approaches in section 1.5.

The research questions and hypotheses tested are stated in section 1.7 after a discussion of teachers' conditional reasoning ability in section 1.6.

1.1 General Goals of the Study

1.1.1 The role of logical analysis in general education.

One important and well recognized goal of any educational system, is an intellectual independence on the part of its graduates. A major obstacle to intellectual independence is the acceptance of hypotheses without giving them critical thought, testing them against alternatives, or checking their consistency with known facts -- in general, without considering them critically. In fact, one cannot be intellectually independent without being able to judge critically one's own and others' deeds and sayings. This is particularly true in the scientific domain. The history of science and mathematics shows several cases where arguments were mistakenly accepted because a valid alternative was overlooked. In everyday life too, there are many examples where carelessly drawn conclusions lead to a decision that later on proves to be wrong. Had the conclusions been double checked for their validity, the wrong decision and possible disappointment could have been avoided. O'Brien, whose research influenced this present study, stated:

> "One aspect of critical thinking is the ability to test the logical validity of an argument, an important ability in everyday life when arguments such as 'Communists favor U.S. withdrawal from Viet Nam. Bill favors U.S. withdrawal from Viet Nam. Therefore Bill is a Communist' are widely accepted as valid." (O'Brien et al. 1971).

Independence in thinking, in short, requires the ability and the knowledge of proof-finding. Because a deductive proof is, by definition, a sequence of inferences made by the rules of logic from previously established facts (theorems), or accepted basic assumptions (axioms), mastery of the basic rules of logic, which enter any proof-finding activity must be a part of the repertoire of any intellectually independent person. The examination of how such a mastery is accomplished is basic to this study. Is logical analysis a learned ability or a developmental one? Does logic emerge spontaneously and achieve its highest degree through the process of growth? Or, does it take a deliberate intervention to improve on it? Can we afford to wait for the appearance of logic in the same way we wait for the appearance of permanent teeth? Different psychologists answer these questions in different ways.

O'Brien and Shapiro (1968) used Hill's (1961) "yes/no" test, modified to a "yes/no/not-enough-clues" test, to study students' ability to discriminate between a necessary conclusion and one that does not necessarily follow from the premises. When children six to eight years of age were called upon to test the logical necessity of a conclusion, they experienced great difficulty. They were rarely able to perform above the chance level when no logically necessary conclusion existed. Further, growth in this ability over the three-year span studied was negligible.

In a later study (1970), O'Brien tested a sample of upper-middle-class children ages six through thirteen in Ohio. He found, in accordance with Hill, that in recognizing logical necessity, i.e., in the ability to apply 'Modus Ponendo Ponens' and 'Modus Tollendo Tollens,' subjects had little difficulty, this ability leveling off at six to eight years of age. However, in testing for logical necessity, i.e., avoiding the fallacies known as 'Affirming the Consequent' and 'Denying the Antecedent,' students exhibited considerable difficulty*. Although there was improvement in test scores over the eight year period, there was no evidence of any leveling off. It becomes apparent from O'Brien and Shapiro's studies that about 75 percent of middle-class elementary-school children consistently misinterpret 'if-then' as 'if-and-only-if,' and as a result, draw invalid conclusions. This misinterpretation was still found to be widespread at the college level and, in a few cases, even in the face of a college-level course in logic (O'Brien 1973).** Chen's cross-age study (1975) indicated that regardless of their ability to recognize that certain conclusions do not follow from the premises, students of various ages seldom manifest this ability. His subjects rarely based their decisions in scientific, social, and judicial problems on careful analysis of the information given unless they were specifically called upon to do so. A speculative synthesis of the data, rather than logical analysis, was found to be the most common response in fifth grade as well as in eighth and eleventh grades.

The fact that O'Brien's university students and Chen's senior-high-school students did not consistently apply the rules of logic in a variety of contexts, indicates on one hand that logic is not a pure developmental phenomenon; at least logic's full development and habitual use are not reached through growth alone. On the other hand, existing school curricula are not providing effective training in logic. Moreover, Eisenberg and McGinty (1974) found that in some specific forms of sentential logic "maturation - and, unseparably, education - seem to have a negative effect on one's ability."

---

*For a broader discussion of the rules of inference mentioned, refer to section 1.3.6 page 18 or to table 1.2 page 20.

**For further discussion of patterns of wrong inferences see sections 1.3.6 (page 21) and 1.7.3 (page 55). See also comments on English interpretations of conditional sentences at the end of section 1.3.3 (page 14).

Some psychological studies consider the influence of differential training on the number of correct conclusions drawn by college-level students in syllogistic reasoning (Wason 1964; Johnson 1966; Pezzoli and Fraze 1968). These studies found that such training reduced the number of incorrect answers given by college students. Is there an effective way to train younger students in proper logical reasoning?

Suppes (1965) took a pure syntactical approach in his one-of-a-kind study of very young children's proof-construction behavior. He devised a content-free task, in which he gave first graders strings of 0's and 1's printed on cards, and asked them to reproduce a string using four reproducing rules (serving as means of inference) and the symbol 1 as a starting point (serving as the simple single axiom). Subjects were divided into two groups. In the "correction group," subjects were corrected for each wrong step in each proof; in the other group, subjects were stopped only when a valid proof was not completed in three times the length of a shortest proof. The correction group was the more successful group; the intervention seemed to speed up the learning process.

A one-shot training in a content-free task is too artificial to be considered for application as part of a school curriculum. To learn anything about children's potential for realistic deductive reasoning, one must involve them in an autonomous, continuous undertaking of deductions involving meaningful content. Children must be allowed to have the opportunity not only to exhibit what they are able to do, but also to exhibit what they are capable of achieving through systematic learning. Exhibiting deductive reasoning through the finding of proofs requires the ability to point out which premises imply a given conclusion and by which rule of logic. This ability in turn presupposes the ability to draw valid conclusions and to avoid fallacies. Systematic introduction of young children to valid and invalid rules of inference was a major goal of this study.

1.1.2 The role of logical analysis in mathematics and in mathematics education. The logic of implication is widely regarded as being at the heart of mathematics, because implications are the links that hold mathematical ideas together through axiomatic methods. Every mathematician spends a considerable part of his time following proofs established by colleagues and looking for valid, clear, elegant proofs for his own mathematical conjectures. This is so because proofs, based on the logic of implication, function in mathematics as a decisive means of persuasion of the truth of a given statement. If one of the reasons mathematics is taught in school is to make children understand how mathematics functions as a part of human culture, the notion of a proof, composed of a sequence of small deductive steps, should be presented at least on the same level of importance as the other strands of mathematical activity, such as computation and application.

While mathematical studies involve the interrelation of several types of activities -- computation, application, abstraction and deduction -- traditionally, K-12 mathematics curriculum has been almost entirely restricted to building various kinds of computational techniques, with a limited amount of application. Deduction has been almost entirely neglected except in high-school geometry. Most high-school mathematics teachers share the experience that students, when first introduced to deduction through proof-finding activity in geometry, encounter tremendous difficulties in understanding the need for a proof and in grasping the concept of a valid proof. One reason for this might be a lack of sufficient earlier preparation that would gradually and continuously train students in finding proofs.

The math curriculum reform of the 1960's attempted to explain elementary computational algorithms by calling attendion to structural elements of the underlying number system. This opportunity for introducing young pupils to deductive mathematics was largely wasted, at least at the elementary level, through an inability of most elementary teachers to deal with the level of deduction needed. Consequently, facts about structure were simply committed to memory, the same way computation facts were previously handled.

Problem-solving activities in school mathematics also fail to take advantage of the occasions for deductions inherent in these activities. They are focused, at all levels, on finding a correct solution rather than justifying the process by which the solution is found. The student is required to work on a problem toward a solution. Rarely is he asked to find out or explain why his methods are correct. Here is a typical example from a fourth grade modern math book: "What are the elements of the intersection of A = {1,2,3,4,5} and B = {1,3,5,7,9}?" As soon as the student comes up with AB = {1,3,5} as a solution, he turns to a new problem. Questions like "Why did you put 1 there?", "Why didn't you put 2 there?", "Is your set the only possible solution?", "Why?" are almost always omitted. Shouldn't a student be able, and be expected, to verify such an answer, particularly if this verification could be accomplished by simply referring back to the definition (most often a conditional sentence) given a few rows above the problem solved? After all, it is only through deductive argument that one can claim that the premises provide conclusive evidence for the conclusion.

In today's math textboods there are a few such natural opportunities for simple deductions, yet they are rarely emphasized. The teachers seem to overlook them altogether or regard them as having minor importance. There is a lack of concerted effort to improve the reasoning ability of elementary school children and accelerate the processes of change and development that occur during the first six school years.

Based on the recognition that deduction plays a very important role in mathematics, and on the desire to introduce children to logical analysis to aid them in their general education, the present study was designed to enrich the school mathematics curriculum in order to redress the distorted view of mathematics created by the current imbalanced emphasis on computational rules. The goal of this study was threefold: to devise a unit containing a significant deductive component comprised of problems that are susceptible to solution by inferential techniques; to have teachers implement it; and to evaluate that unit by measuring the induced change in learners' ability to apply valid rules of inference and to avoid fallacies. As Kyrgowka (1971) pointed out,

> "the essential problem is to work out educational approaches to (a) initiation of the pupil into the process of active axiomatic thinking (axiomatization, deduction, interpretation) adapted to the level of his intelligence and the content of the curriculum; (b) the formation in the pupil's mind of a concept which, although not yet formalized, would nevertheless correspond in its essentials to the modern axiomatic concept."

1.2 Specific Goals - Conditional Reasoning

Mathematics is developed by inferences made from axioms, theorems, and definitions. Every theorem in any mathematical subject is, in effect, a conditional sentence; any definition is a biconditional one. Knowing how to make valid inferences from conditional statements and how to avoid fallacious reasoning from these statements is therefore essential for any consistent development of mathematics.

The present study attempts to contribute to the achievement of previously discussed general goals, i.e., education that leads to intellectual independence and integrative school mathematics curricula, by developing a unit that will enable students to apply to conditional statements two valid rules of inference, 'Modus Ponendo Ponens' and 'Modus Tollendo Tollens,' and to avoid two fallacious rules of inference, 'Affirming the Consequent' and

'Denying the Antecedent.'*

The selection of these four rules was supported by previous studies that provided data about the development of the ability to apply the two valid rules, and offered explanations for errors usually committed when the two nonvalid rules were incorrectly applied.

Hill (1961) had 100 items in her test, of which 51 included conditional premises of sentential logic. She rank-ordered the difficulty of each logical form by average number of errors per test item. In ranks 1 to 11, there are 18 items of which 13 are from these 51 items, an additional 2 belong to sentential logic but do not have conditional premises, and only 3 are non-conditional items (of quantificational logic). These findings too justify special attention to reasoning with conditional premises.

Robert Kane (1975) discussed the proof-making task as a complex-terminal behavior. Associated with this task is a set of prerequisite or subsidiary behavior. He favors approaching the teaching of proof construction by systematically teaching the prerequisites one by one, and then combining one with another. Kane was unable to offer a hierarchy of prerequisites. Nevertheless, he discussed bits and pieces of its structure. Among those he mentioned:

> "Give examples of the use of Modus Ponens; Give examples of converse of implications; Use the fact that $(-q \rightarrow -p) \leftrightarrow (p \rightarrow q)$ to set up a strategy for proof by contrapositive; ...; show by a counter example that $p \rightarrow q$ and $q \rightarrow p$ are not equivalent."

---

*Specification of the underlying logic is given in section 1.3, page 11.

These prerequisites, which Kane, like most mathematicians, considers essential to the learning of proof-finding, are closely related to conditional reasoning, and support the need for a unit of school instruction on the rules of inference with which the present study was concerned.

1.3 Mathematical Logic Underlying Conditional Reasoning

In this section the mathematical model of the patterns of deduction relevant to the present study, as well as related concepts of symbolic logic, are laid out in detail. It is intended for the non-professional logician. For whoever may wish to skip the details, a brief summary is given in tables 1.1 (page 14), 1.2 (page 20), 1.3 (page 24), and 1.4 (page 26). A more elementary account can also be found in Appendix 7.1 (page 209).

1.3.1 The language of sentential conditional logic. Let p_i, q_i $i=1,2,3...$ be symbols for distinct sentences. Let "-" be the negation symbol, interpreted in English as "not," and let "→" be the conditional symbol, interpreted in English as "if...then...". We assume that none of these symbols is a finite sequence of the other symbols. In particular the previous statement means that no sentence symbol is a combination of any other sentence (or other) symbols e.g., $p_1 \neq -p_2$ and $p_{14} \neq q_7 \rightarrow p_{53}$.

Any finite sequence of the above symbols is a *formula*. Among all possible formulas we single out the "grammatically correct" ones by specifying what a *well formed formula* is:

(i) Every sentence symbol is a wff (well formed formula). These wffs will be referred to as atomic wffs.

(ii) If a,b are wffs, then -a (not a); a → b (if a, then b), are wffs.

(iii) No other formula is a wff.

One should bear in mind that the purpose of this discussion is to build an abstract model of the logic of conditional reasoning as expressed in English. The wffs then are abstract objects interpreted into English as simple declarative sentences, their negatives, or conditional sentences, built up from the declarative sentences.*

For example: Let p stand for: "Mary is sick," -p will then be "Mary is not sick." Let q stand for "Mary goes to school," -q will then be "Mary does not go to school." We can now obtain sixteen new grammatically correct sentences by introducing the if...then... connective to tie any two of these sentences into a conditional sentence. For example: p → q, which means "If Mary is sick, then Mary goes to school," is a grammatically correct sentence even though it does not make too much sense. Also; -p → -p, which means "If Mary is not sick, then Mary is not sick," is all right from the grammatical point of view even though it does not reveal any new information. The process of building up new wffs from ones previously obtained may go on and on and take an infinite number of paths. Some wffs will make more sense than others in certain English interpretations. However, it is not the content meaning that symbolic logic depicts but rather the grammatical structure of the language. Logic is the study of content-independent language forms that constitute valid arguments.

---

*For a complete account of the language of sentential logic see Church, 1956, chapter 2.

1.3.2 Conditional sentences and their negation modes.

This study considers only conditional wffs with one occurrence of the conditional connective, and with, at most, one occurrence of negation on each side of the conditional. In other words, sentences such as the following will not be considered, even though formally they are wffs and whatever will be said holds for them also: "If, 'If Mary is sick then she does not go to school,' then 'If Mary goes to school then she is not sick.'"

Back to the simple case. A simple conditional sentence* may assume one of the following four negation modes, depending on the location of negations: $p \rightarrow q$; $p \rightarrow -q$; $-p \rightarrow q$; $-p \rightarrow -q$. These modes will be denoted by ++; +-; -+; --; respectively. E.g., the first example given in section 1.3.1 is in ++ negation mode, and the second one is in -- negation mode; $a \rightarrow b$ will denote a simple conditional sentence in any negation mode, where a,b may be either atomic wffs or their negations.

Any conditional sentence $a \rightarrow b$ can be partitioned into two parts. The first part, a, is called the antecedent, the second part, b, is called the consequent.

1.3.3 Conditional sentences - their truth value and English interpretation.

Considering interpretations of a given wff in ordinary English, any interpretation of a sentential wff will be either a true (declarative) sentence or a false one. The truth or falsity of a conditional sentence depends on the truth or falsity of its antecedent and consequent in the following way:

---

*The words "sentence" and "wff" are used interchangeably, and not in the sense of a formula with no free variable.

Table 1.1 Truth Value of Conditional Sentences
(T = true, F = false)

If the truth values of a,b are:		then the truth value of a → b is:
a	b	a → b
T	T	T
T	F	F
F	T	T
F	F	T

This truth table is justified mainly because the only case that contradicts a → b is when a holds but b does not; any other combination of truth values for a and b does not disprove a → b and, therefore, does not prevent a → b from being a true assertion. Although this truth-value analysis does not represent the full complexity of use of "if...then..." in idiomatic English, it describes fully and precisely the use in mathematical language. (For more details see Quine, 1951, page 14).

As Suppes (1957) points out, several other idioms in English have approximately the same systematic meaning as "if...then...". For example a → b also means: a only if b; b if a; b provided that a; a is a sufficient condition for b (the occurrence of the antecedent suffices to guarantee the occurrence of the consequent); b is a necessary condition for a (without the consequent occurring, the antecedent does not occur). It is a common "mistake" to use "if" in the sense of "only if" and vice versa. For example when a mother says to a child in the morning: If it rains after school, call me to pick you up" she probably means "...otherwise you should walk home." In other words, she used "if it rains etc." to mean "only if it rains etc." Also, many times changing the expression from the "if...then..." style to the "only if" style makes much more sense, particularly when the content implies causality. For example: "If it rains tomorrow, then it will be cloudy tomorrow" makes less sense than the equivalent expression: "Only if it is cloudy, will it be rainy tomorrow." This is because one usually does not look for clouds when it rains, but one does look for a clear sky (i.e., not cloudy), to make sure it will not rain.

1.3.4 Sentences equivalent to conditional sentences. Two sentences are said to be tautologically equivalent if and only if they have the same truth table. As can be checked by reference to the truth tables of the sentential connectives "and," "or," "not" (denoted $\wedge$, $\vee$, $-$, respectively) to be found in any introductory logic textbook (e.g., Mendelson, 1964), each of the following sentences is tautologically equivalent to any of the others:

(i) $a \rightarrow b$

(ii) $-(a \wedge -b)$

(iii) $-a \vee b$

(iv) $(a \wedge b) \vee (-a \wedge b) \vee (-a \wedge -b)$.

To give one example for each:

(i) If it is 9:00 a.m., then the bell rings.*

---

*This example includes a free variable, "it," and in fact its analysis falls in the category of first order logic. See section 1.3.7, page 21.

(ii) It never happens that at 9:00 a.m. the bell does not ring.

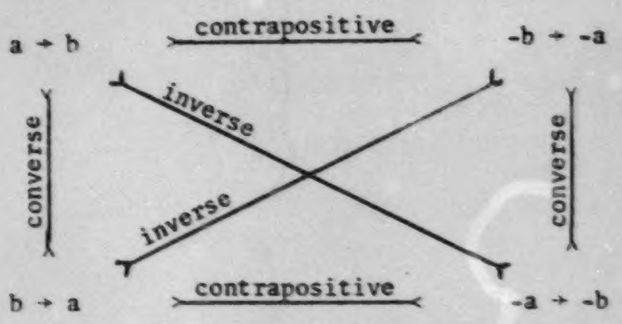
(iii) Either it is not 9:00 a.m. or the bell rings.

(iv) Either it is 9:00 a.m. and the bell rings, or it is not 9:00 a.m. and the bell rings, or it is not 9:00 a.m. and the bell does not ring. (The fourth alternative is eliminated by (ii).)

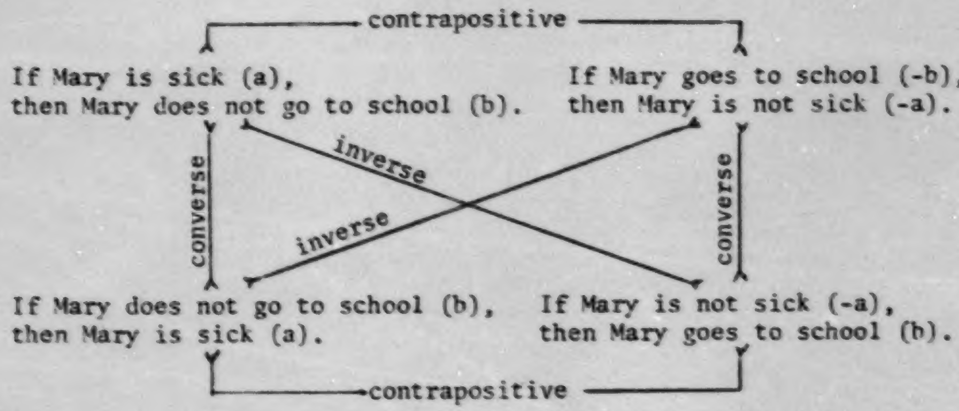
1.3.5 Relations betweed conditional sentences. Two conditional sentences in which the antecedent and the consequent are interchanged are said to be converse to each other. For example the sentence "If Mary is sick, then Mary does not go to school" and the sentence "If Mary does not go to school, then she is sick," are converse to each other. Notice that in this example the first one is in the +- negation mode, and the second is in the -+ negation mode, because the two parts were interchanged. The converse of a ++ sentence will remain a ++ sentence, and similarly for a -- sentence. Notice also that a conditional sentence and its converse do not reveal the same information. In the above example even if the first sentence is true, the second one does not have to be true, and it certainly does not follow from the first one.

There are other relations between pairs of conditional sentences. The structural connections are illustrated in the following diagrams:

Diagram 1.1 Relations between pairs of conditional sentences


a → b
contrapositive
-b → -a
converse
inverse
inverse
converse
b → a
contrapositive
-a → -b

For example:



If Mary is sick (a),
then Mary does not go to school (b).

If Mary goes to school (-b),
then Mary is not sick (-a).

If Mary does not go to school (b),
then Mary is sick (a).

If Mary is not sick (-a),
then Mary goes to school (b).

contrapositive
inverse
converse
inverse
converse
contrapositive

Notice that because, in this example, b includes a negation, -b does not, for in English a double negation is replaced by a positive expression. Again we have a → b in +- negation mode, and therefore its converse and inverse are in -+ mode and its contrapositive is back in +- mode.

From a given conditional sentence a → b, its contrapositive, -b → -a, follows. In fact, two contrapositive sentences are logically equivalent. They express identical content and hence they are both either true at the same time or false at the same time, but never of opposite truth value.

1.3.6 Two valid and two nonvalid rules of inference. A valid rule of inference enables one to deduce from given premises a new sentence, regardless of the truth or falsity of the premises. The conclusion will certainly be true if the premises are true. Nonvalid reasoning process may sometimes yield true conclusions, but often it yields false ones, even from true premises.

The fundamental rule of inference is called Modus Ponendo Ponens (or the law of detachment) abbreviated as MP. It is the rule by which one infers b from the pair of sentences: $a, a \rightarrow b$, where a,b are wffs as mentioned in section 1.3.1.

Similar to MP is the rule of inference known by the Latin name Modus Tollendo Tollens abbreviated as MT. It allows the derivation of -a from the pair of sentences: $-b, a \rightarrow b$. This rule is sometimes referred to as the law of the contrapositive.

The following are two examples intended to illustrate the validity of conclusions drawn by applying the above rules of inference. (The use of nonsensical terms is taken from Enderton 1972; the terms themselves are from the classic poem by Lewis Carroll.)

From: If (a) it is brillig, then (b) borogoves are mimsy.
and from: (a) It is brillig.
Conclude by MP that: (b) borogoves are mimsy.

Note that we can make this inference without the slightest idea of what a mimsy borogove is. Similarly,

from: If (a) it is brillig, then (b) borogoves are mimsy.
and from: (-b) Borogoves are not mimsy.
Conclude by MT that: (-a) it is not brillig.

Because if it was brillig, then by the first premise borogoves would have been mimsy -- which then would contradict the second premise. Such a contradiction is unbearable by the law of excluded middle. (For a discussion of the principles of proof by contradiction see Suppes, 1957, page 38.)

A second look at the two rules MP, MT will reveal that they are both patterns of inferences from a conditional sentence a → b and one other sentence -- either an affirmation of its antecedent in the case of MP (which yields an affirmation of the consequent), or, in the MT case, a denial of its consequent -b, (which yields the denial of the antecedent.)

Suppose now the premises are a conditional sentence (a → b) and an affirmation of its consequent (b). Then it would be nonvalid to infer either the affirmation of the antecedent (a) or its denial (-a). If an inference like that is made it is called the fallacy of Affirming the Consequent, denoted by AC.

Similarly, it is nonvalid to infer either the denial (-b) of the consequent or the consequent (b) itself from a conditional sentence (a → b) together with the denial of its antecedent (-a). Doing either is referred to in logic as applying the fallacious rule called Denying the Antecedent, denoted DA. Table 1.2 summarizes the rules (page 20).

It should be noted that if one mistakenly takes a → b to imply its converse b → a, then affirmation of the consequent (b) turns out to be an affirmation of the antecedent (of the converse). Thus the nonvalid derivation of a in this case may be explained by application of the valid MP to the converse conditional sentence assumed mistakenly to hold. Similarly, the fallacious derivation

Table 1.2 Rules of Inference and their Validity

Premises	Conclusion(s)	Name of rule of Inference (abbreviated)	Valid/ Nonvalid	Example
1. $\begin{cases} a \to b \\ a \end{cases}$	b	Modus Ponendo Ponens (MP)	valid	If (a) Jane is Jack's sister, then (b) she* lives on Washington St. (a) Jane is Jack's sister. ∴ Jane lives on Washington St. (b).
2. $\begin{cases} a \to b \\ -b \end{cases}$	-a	Modus Tollendo Tollens (MT)	valid	If (a) Jane is Jack's sister, then (b) she lives on Washington St. (-b) Jane does not live on Washington St. ∴ Jane is not Jack's sister (-a).
3. $\begin{cases} a \to b \\ b \end{cases}$	a, -a	Affirming the Consequent (AC)	nonvalid	If (a) Jane is Jack's sister, then (b) she lives on Washington St. (b) Jane lives on Washington St. ∴ Jane may or may not be Jack's sister.
4. $\begin{cases} a \to b \\ -a \end{cases}$	b, -b	Denying the Antecedent (DA)	nonvalid	If (a) Jane is Jack's sister, then (b) she lives on Washington St. (-a) Jane is not Jack's sister. ∴ Jane may or may not live on Washington St.

*The word "she" is in fact not a free variable in the sentence b because in the natural use it refers to Jane. This is why in the conclusion her name appears. For further discussion see note on free variables in section 1.3.7.

of not-b from the pair: a → b and not-a, may be explained as an application of MT to the converse b → a and not-a (or the application of MP to the inverse not-a → not-b), which then will validly yield not-b. This tendency to consider the converse (or the inverse) in place of a given conditional and to draw a definite conclusion (thus committing an AC or DA error) is systematic in most cases, and explains another name common to these two fallacies: "Assuming the converse " - for AC; and "assuming the inverse" - for DA.*

1.3.7 Extending the relevant rules of inference to quantificational logic.

Sentential logic is a very limited model of deductive reasoning. There are many examples of intuitively correct deductions that cannot be adequately mirrored in that model. The case most relevant to this study is exemplified by the following: "If a positive integer has more than two divisors, then it is not a prime number. 26 is a positive integer which has more than two divisors. Therefore 26 is not a prime number." Even though this deduction very closely resembles the MP pattern of inference, in fact it is not a pure application of MP. First of all, the second given sentence is not exactly the antecedent of the given conditional sentence, but is rather an instantiation of the antecedent. Second, the conditional sentence itself is not really one that can be represented by a wff in sentential logic, because it refers to an arbitrary (unspecified) number. Its antecedent "a positive integer has more than two divisors," is not really a sentence and does not have a well-defined truth value. The antecedent is neither a true nor a false assertion since the expression "a positive integer" allows for a wide range of numbers, some of which have more than two divisors, others exactly two, still others fewer than two. Therefore no wff of sentential logic can take the antecedent as an interpretation.

---

*For further discussion of errors in conditional reasoning see section 1.7.3; also see Stabler, 1953, page 73.

"A positive integer" is an example of a free variable, i.e., a variable which is uncontrolled by a quantifier. To remedy the antecedent of the above conditional sentence one could say, for example: "There exists a positive integer which has more than two divisors." The preceding is a true sentence. Or, one could say: "All positive integers have more than two divisors," which is a false sentence. But these remediations change the assertion expressed in the conditional sentence. The conditional sentence "if a positive integer has more than two divisors, then it is not a prime" says in fact: "For all x, if x is a positive number with more than two divisors, then x is not a prime." As such it belongs to first-order (or predicated, or quantificational) logic. In this modified sentence x is no longer a free variable because it is controlled by the quantifier "for all." The truth or falsity of such a conditional sentence is no longer determined by the truth or falsity of its antecedent and its consequent alone but rather by a check of all possible interpretations of this sentence, namely by putting every positive integer to the tests of having more than two divisors, and of being a prime. Because no positive integer that has more than two divisors is a prime, this sentence is true, and the deduction with the number 26 based on this sentence is valid.

Expressions of natural English many times are inaccurate in the sense exhibited in the above example; namely they seem to include a free variable but in fact there is an implicit quantifier controlling it, which is not explicitly mentioned. This, in fact, happened in our example twice, once in the antecedent as discussed, and again in the consequent when the word "it" was used.

In this study, examples like the above appear very often, and usually they are not distinguished from pure sentential ones. For the sake of precision, counterparts of the sentential inferential forms MP, MT, AC, and DA are given in table 1.3, page 24.

1.3.8 Algorithmic solutions to validity judgments.

In this study the learner in most cases was confronted with a situation, was given an inference drawn from premises based upon that situation, and was asked to judge its validity. This judgment was based on reasoning within the context in which the deduction was made. Arguments always require inference from the premises' meaning (semantics).

For example,* a device with eight switches, each turning on and off one of eight possible combinations of three colored light bulbs, was given to a group of students to play with. After a while students expressed facts about the switch board such as: "If switch number 1 is pushed, then the yellow light is on." The board was then covered. Teacher pushed a switch and uncovered the yellow bulb only. Suppose it was on. Teacher would then ask: "Did I push switch number 1?" The right answer is "you may or may not have done so because switch 1 is one of many switches that turn on the yellow light (possibly with other lights)." This

---

*For more examples see section 2.1 (page 58) and appendix 7.1.

Table 1.3 Rules of Inference Extended

x denotes a variable; a,b are one-place predicate symbols;
t is a constant symbol; and "$\forall$" means for all.*

Premises	Conclusion	Rule of Inference	Example
$\forall x[a(x)\rightarrow b(x)]$ $a(t)$	$b(t)$ valid	MP	If [a(x)] someone plays too much football, then [b(x)] he does not do enough homework.** John plays too much football [a(t)]. ∴ John does not do enough homework [b(t)].
$\forall x[a(x)\rightarrow b(x)]$ $-b(t)$	$-a(t)$ valid	MT	If [a(x)] someone plays too much football, then [b(x)] he does not do enough homework. John does enough homework [-b(t)]. ∴ John does not play too much football [-a(t)].
$\forall x[a(x)\rightarrow b(x)]$ $b(t)$	$a(t);-a(t)$ non-valid	AC	If [a(x)] someone plays too much football, then [b(x)] he does not do enough homework. John does not do enough homework [b(t)]. ∴ John may or may not play too much football.
$\forall x[a(x)\rightarrow b(x)]$ $-a(t)$	$-b(t);b(t)$ non-valid	DA	If [a(x)] someone plays too much football, then [b(x)] he does not do enough homework. John does not play too much football [-a(t)]. ∴ John may or may not do enough homework.

---

*The existential quantifier is not introduced because it rarely appears relevant to this study.

**Here we again have an implicit quantifier, for in fact this sentence conveys the following information: for all students x, if x plays too much football [a(x)], then x does not do enough homework [b(x)].

answer was abbreviated to "not enough clues" followed by an argument similar to the one given above. Syntactical arguments were never acceptable as answers given by the students and were certainly never presented by the teacher or explicitly mentioned. Their existence, however, was suggested by the inductive nature of the presentation.

Table 1.4 (page 26) gives the algorithm which, if slavishly followed, gives the right answer to any validity judgment relevant to this study, regardless of meaning and context. For simplicity, only the sentential model appears. The table for the first-order model is similarly obtainable.

1.4 Psychological Research Underlying Conditional Reasoning

1.4.1 The relations between symbolic logic and the psychological reasoning process. In section 1.3 some logical forms of valid and nonvalid deductions from conditional premises were discussed. However, symbolic logic is not intended to be, and indeed is not, a model of the reasoning processes going on in the mind while deductive thought takes place. Logic provides objective criteria for judgment of the validity of the outcome of the reasoning process. To merit a favorable judgment, one must arrange his arguments in a sequence, and check the inferences against the criteria provided by logic. But this sequence has only rarely any similarity to the temporal sequence of thought by which one reaches these conclusions.

In the particular case of the test involved in the present study (see appendix 7.2, page 321) almost no thinking is needed to

Table 1.4 Algorithmic Solution to Relevant Problems of this Study

p,q are distinct simple (atomic) sentence symbols, in which no negation occurs;

$\to$ is the conditional symbol interpreted "if...then...";

$-$ is the negation symbol;

$p?$ $q?$ are the interrogative sentences corresponding to p,q;

"clues" is the term used in this study for the given information on which one should base the answer;

"not enough clues" is an abbreviated term for the idea that the information given in the clues does not suffice to yield a definite "yes" or "no" answer to the question;

NOTE: The question never includes negation.

Negation Mode \ Logical Form		MP	MT	AC	DA
++	clues:	$p\to q$ p	$p\to q$ $-q$	$p\to q$ q	$p\to q$ $-p$
	question:	$q?$	$p?$	$p?$	$q?$
	answer:	yes	no	Not enough clues (NEC)	NEC
+-	clues:	$p\to -q$ p	$p\to -q$ q	$p\to -q$ $-q$	$p\to -q$ $-p$
	question:	$q?$	$p?$	$p?$	$q?$
	answer:	no	no	NEC	NEC
-+	clues:	$-p\to q$ $-p$	$-p\to q$ $-q$	$-p\to q$ q	$-p\to q$ p
	question:	$q?$	$p?$	$p?$	$q?$
	answer:	yes	yes	NEC	NEC
--	clues:	$-p\to -q$ $-p$	$-p\to -q$ q	$-p\to -q$ $-q$	$-p\to -q$ p
	question:	$q?$	$p?$	$p?$	$q?$
	answer:	no	yes	NEC	NEC

answer the test items once the algorithmic solution provided by symbolic logic has been mastered. So, the purely syntactical inferences established by symbolic logic should not be confused with the psychological thinking processes of human reasoning, which usually involves semantic and other considerations.

The use of symbolic logic as an ideal model for psychological processes is widespread in literature (e.g., Piagetian description of the stage of formal thought using the 16 binary operations, see section 1.4.2). That symbolic logic should not be used as a model of thinking processes is true not only because application of formal logic sometimes "saves" thinking, but also in view of the many interfering psychological factors that jeopardize logical analysis. Some factors are related to the emotional impact of the context. When children are involved in logical analysis, their stage of intellectual development and level of language attainment may also have an effect on their ability to reason logically. Still another influential factor is previous experience through interaction with adults like parents and teachers. Sections 1.4 and 1.5 discuss some of the psychological research done to shed light on these problems, particularly with respect to conditional reasoning.

1.4.2 Readiness. Granted the indubitable need for a deliberate effort to bring about full awareness of the difference between conclusions that necessarily follow from given premises and those that do not, it is natural to seek indications of the right age at which to exert substantial effort in this direction.

As early as 1919 the British psychologist Burt claimed, in light of his research results, that all the elementary mental mechanisms essential to formal reasoning are present by the mental age of seven. A child's reasoning ability appears to be a function of the degree of organic complexity of which his attention is capable. Dorothy Wheeler's experimental work (1958) supports Burt's results. These psychologists' findings disagree with the developmental theory of J. Piaget, who holds that mental growth proceeds by stages.

According to Piagetian theory and research, the ability to reason logically marks the formal thought as the highest developmental stage, appearing only at adolescence. Piaget uses symbolic logic to refer to reasoning processes. A major component of formal thinking is the sixteen binary operations. Take any two statements p and q and their negations $-p$, $-q$. Using the "and" connective denoted by $\wedge$, one gets four new statements: (i) $p \wedge q$; (ii) $-p \wedge q$; (iii) $p \wedge -q$; (iv) $-p \wedge -q$. There are 2^4, that is 16, ways to invent still new sentences from these four using the "or" connective, denoted by $\vee$. With respect to the present study, "if p, then q" which is logically equivalent to $(p \wedge q) \vee (-p \wedge q) \vee (-p \wedge -q)$, is one of the sixteen. It is the ability to formulate and apply any of the sixteen combinations, which Piaget refers to as the major component of formal thinking. In particular,

> "The role of possibility is indispensible to hypothetico-deductive or formal thinking...the connection indicated by the words "if...then" links a required logical consequent to an assertion whose truth is merely a possibility" (Inhelder and Piaget, 1958).

Piaget (1959) showed that it is difficult for children at earlier stages, more precisely at the concrete-operation stage, to accept a hypothetical assumption and draw conclusions from it. Research done by Shirley Hill (1961) provided evidence that middle-class Californian children at ages six, seven, and eight are already able, with a high degree of success, to recognize valid conclusions from hypothetical premises (where validity is determined by formal logical principles). O'Brien and Shapiro's results mentioned in section 1.1.1 brought refinements to Hill's observations. In a cross-age study they found students confusing if-then statements with if-and-only-if statements. This confusion is determined by deriving the antecedent in AC problems, and deriving the denial of the consequent in DA problems, instead of realizing that not enough information is available to make any derivations in these cases.* "Whether children's hypothetical-deductive abilities can be altered by some systematic intervention, is an open question" say these researchers (1970).

J. Roberge (1970) was concerned with differential development with respect to six specific principles of deductive reasoning in class and conditional reasoning. He found that fallacies are the most difficult to analyze at each grade level tested: fourth, sixth, eighth, and tenth grades at three suburban schools in Connecticut. The logical forms AC and DA were not mastered by any substantial percentage of students prior to the tenth grade. The mastery of MP reached 95 percent level in conditional reasoning. His results suggest, he said,

> "That classroom instruction of the valid principles of class and conditional reasoning, such as MP, could begin as early as fourth grade."

Lester (1975) in following Suppes' (1965) idea of using strings of 1's and 0's to study children's proof construction behavior (see section 1.1.1), found that

> "subjects in the upper elementary grades (4-6) are able to solve problems in this system as successfully as the older subjects, except that they require more time...

---

*Refer to section 1.3.6 for interpretation of AC and DA (page 18).

> there is reason to believe that even students in the upper elementary grades can be successful at mathematical activities that are closely related to proof. That is, certain aspects of mathematical proof can be understood by children nine years old or younger."

These studies encouraged the present investigator to start working with fourth and fifth graders.

1.4.3 Content effects. "Discerning observations and experimental evidence must lead one to conclude that rational thinking is not free from the influence of the affective processes," says A. Lefford (1946) in a study that demonstrated the effect of verbal stereotypes on syllogistic reasoning. He found that most subjects solve neutrally toned syllogisms more easily than syllogisms with controversial matters which were likely to arouse some affective reaction to their conceptual subjects. For a long period most psychologists tended to equate logical thinking with syllogistic reasoning.

As early as 1928 Minna Wilkins scrutinized the effect of a syllogism's content on the ability of one who considers this syllogism to accept valid conclusions. She found familiar material to be the least disturbing to logical thinking and material of unfamiliar scientific or nonsense terms to be the most deleterious, with suggestive and symbolic material in between.

When subject matter was likely to play on emotions, prejudices, or attitudes, the reasoning process was often distorted by Feather's subjects (1964). His results show that when subjects agreed with the conclusion they made more errors in accepting invalid conclusions than in rejecting valid ones. When they disagreed with the conclusions, more errors were made in rejecting valid conclusions than in accepting invalid ones.

Parrott (1967) studied the effects of premise content on accuracy and solution time in syllogistic reasoning. He found solution time and number of correct responses to be positively correlated with each other on true, false, and mixed premise-content.

Most of these studies used college students as their subjects. It is therefore interesting to refer to a study dealing with critical-thinking readiness in grades 1-12 by Ennis and Paulus (1965). In the first phase of this study they found that on their class-inclusion reasoning test, the concrete-familiar-content component was in general easier for adolescents than the symbolic presentation and the suggestive content. However, on the conditional reasoning test, which is most relevant to the present study, the three components were of about equal difficulty at each adolescent grade level.

The following example should demonstrate the problem of content effect. The fact that the proposition: "all trees are blue" is factually false, makes it harder to follow the syllogism: "all trees are blue, all blue things are not green, therefore all trees are not green," although the conclusion is a logically valid consequence of the premises.

1.4.4 Language effect. Syllogism with purely symbolic terms, i.e., content-free syllogisms, have some of the advantages possessed by lists of nonsense syllables in memory experiments: freedom from extraneous associations and from factual truth or falsity. Some aspects of language, regardless of content, were studied through presentation of content-free syllogisms.

When symbolic syllogisms, valid and invalid, were presented in an experiment, they were found to be unequal in difficulty. Why some syllogisms present easy problems and others difficult ones is the question to which the authors of the following five studies addressed themselves, in an attempt to formulate and verify hypothetical answers.

Woodworth and Sells' (1935) was a preliminary study for Sells' (1936) study. Both studies were concerned with verifying the "atmosphere hypothesis". They considered syllogisms in quantificational logic based on pairs of premises of the following forms: all P are Q (A), all P are not-Q (E); some P are Q (I); some P are not-Q (O).* These syllogisms are called class inclusions in later psychological literature.

Sells found that for 16 possible paired combinations of the four kinds of premises, acceptance of I conclusions always exceeded acceptance of conclusions A. Acceptance of O conclusions exceeded those of E in all but one borderline case, and either I or O was the preferred error for all but one of the sixteen. His formulation of "atmosphere effect" was advanced as accounting for these error preferences. (See table 1.5 for the definition of atmosphere effect.) However the nature of his test format might be expected to dictate high scores on I or O. His tests were constructed of a given syllogism (two premises and a conclusion), and the subject had to put a circle around one of the following four alternatives: AT (if he thought that the conclusion was absolutely true on the basis of the statements), PT (which means probably true), I (indeterminate), AF (absolutely false). Sells counted AT and PT as agreement, and I and AF as disagreement (Ibid., page 60). If an A conclusion is accepted by a student (e.g., all G's are B), logically an I must be also (e.g., some G's are B) unless the set of G's is empty, and similarly if an E conclusion is accepted an O must also be accepted. Therefore, if the subjects were self-consistent on Sells' test, all those subjects who regarded an A or E conclusion as acceptable for a given premise pair would also regard as acceptable the I and O conclusions, respectively, when these were offered because, the empty set case tends to be overlooked. Thus, I or O acceptances should never be smaller in number than those of A or E on this true-false format, and would be expected to be larger. Hence, some of Sell's findings might be an artifact of his 180-items-test format, rather than being attributed to "atmosphere."

---

*The letters A,E,I,O used by these researchers follow a common pattern in logical literature of these days (see Cohen and Nagel 1957).

Chapman and Chapman (1959) re-examined the atmosphere effect, as stated by Woodworth and Sells (1935) and restated by Sells (1936). For their study the Chapmans constructed a syllogism test which consisted of 42 experimental items and 10 filler items, each containing two premises and five alternative conclusions, e.g.,

Some L's are K's.
Some K's are M's.
Therefore:
1. All M's are not L's.
2. Some M's are L's.
3. Some M's are not L's.
4. None of these.
5. All M's are L's.

The correct answer for all 42 experimental items was "none of

Table 1.5 Errors in Syllogistic Reasoning

Atmosphere effect Woodworth and Sells, 1935 Sells, 1936	Conversion and probabilistic reasoning Chapman and Chapman, 1959
Definition: Drawing of conclusions on the basis of global impression of the premises. AA → A; EE → E; II → I, OO → O. Subprinciples: a) a combination of a universal and a particular premise produces a particular atmosphere AI → I; AO → O; EI → O; *EO → O. b) a combination of an affirmative premise with a negative atmosphere AE → E; (AO → O); (*IE → O); IO → O. Principle of caution: a tendency to accept weak and guarded conclusions rather than strong ones (some are, some are-not rather than all are, all are not).	a) Conversion: Interpretation of A and O propositions to mean that the converse is true. This is a result of real experience. (acceptance of the conversion is valid for E and I propositions.) AA → A; AE → E (EA is always valid); IA, AI → I; AO, OA → O *IE → E b) Probabilistic inference like: some P's are M's, some M's are S's, therefore some S's may be P's. II → I. a + b) Combination of invalid conversion and probabilistic inference: OI, IO → O; (IE → E); EE → E; OO → O *OE, EO → either E or O.

*The two models differ on this inference.

these" which was explained in advance to mean: no other alternative among the suggested ones is a valid conclusion of the premises. The five alternative conclusions were assigned randomly to the five positions, with the restriction that each alternative appeared the same number of times in each position. The 10 filler items for which a valid conclusion could be reached were included to prevent subjects from discovering that none of the experimental items had a valid conclusion except "none of these."

An objective comparison of the findings by Sells to those of Chapman and Chapman shows, see Table 1.5, that they agree on predicting the errors of 11 of the 14 possible pairs of premises. (Two pairs, EA and EI, yield valid conclusions and hence cannot be included in any prediction of error.) However Chapman and Chapman suggest a different interpretation for the source of errors. Hence, although they were motivated by criticism of Sells' experimental approach, they ended up with similar results and offered new principles, "conversion error" and "probabalistic reasoning," for obtaining them. This much could probably be done on a theoretical level, and does not require a new experiment with 222 introductory psychology class students. Moreover, their theory would be sounder and the difference in approach between this theory and Sells' would be clearer if the discussion was based on Sells' experimental results with a heterogeneous group of 65 adults.

As mentioned above, atmosphere predictions and the Chapmans' logical predictions by conversion and probabilistic inference are identical for most pairs, differing only with respect to IE, EO, and OE pairs. Begg and Denny (1969) tried to reconcile these differences. The Chapmans, by the principle of probabilistic inference, predicted an E response to an IE pair, but probabilistic inference could equally well predict an O response, congruent with the atmosphere prediction. Since the Chapmans found predominantly E responses, while Sells found O, the question remained an empirical one. In EO and OE pairs the Chapmans predicted and found equal frequency of E and O response, whereas atmosphere effect predicted, and Sells found, O to be the main error response. The purpose of Begg and Denny's study was to gather further data to supply an empirical proof for their claim to reconciliation.

Begg and Denny's study and report indeed completed the theory of atmosphere effect. It was sufficiently descriptive; they used correlation coefficients for statistical comparisons, and obtained highly significant results.

The models of atmosphere effect, conversion errors, and probabilistic reasoning fit into O'Brien's findings of errors in arriving at a definite conclusion in AC and DA items of conditional reasoning, due to considering a conditional statement to be a biconditional one. These models are therefore highly relevant to the present study, even though most of the studies cited in this paragraph used adults or college students for subjects.

1.4.5 Language attainment. Since the present study sought to work with young children, a content-free, pure symbolic approach was not considered, for obvious reasons. However, another aspect of language divorced from content was considered crucial: the minimum level of language attainment necessary for success in undertaking meaningful deductions.

Both Piaget (1959, 1968) and Vigotsky (1965), despite many disagreements between them, quote children at ages nine to ten as using terms like "therefore," "because," "at least," "only," which are typical of the language of deductive arguments, in an acceptable way, namely in their function for logical reasoning. Piaget, who claims that deductive reasoning as an abstract operation reaches its full development no earlier than at age thirteen, does admit that in a concrete situation, a child at age nine to ten is most often at the developmental stage suitable to draw valid conclusions, and to express some of them in a precise way.

The impact of the verbal environment in mathematical classrooms on the logical ability of young children was studied by Gregory (1972). He found that frequency of use of the language of conditional logic by the mathematics teacher changed seventh graders' conditional-logic ability. Later on Gregory and Osborn (1975) studied logical conditional reasoning ability and teachers' verbal behavior within the mathematics classroom. They compared students of teachers who ranked highly in the use of conditional sentences with students of teachers with a low ranking. Their research, they say "has identified the frequency of teacher use of logic as a significant variable in children's acquisition of logic."

All these studies indicate that even if it is true that fourth and fifth grade children sometimes appear to misuse words naturally associated with logical thinking, or may not fully comprehend them, it is reasonable to assume that they possess the basic vocabulary necessary to start expressing logical argument and to begin learning the language for it, under the stimulus of a suitable curriculum that requires logical argument. Because of the growing need for communication, the early years are the best years for language attainment.

1.4.6 Context effects. To avoid the learning difficulties imposed by content on reasoning, experimental psychologists tend to keep their studies content-free. They do so by using symbolic representation, by formulating grammatically correct sentences with nonsense syllables instead of words, or by using unfamiliar terms and content. In the present study these resolutions could not be employed, due to the subjects' young age and the desired educational value of the experimental unit as a vehicle in schools for preparing students gradually for future work with implications within a meaningful and familiar mathematical content. It is largely relevant, therefore, to consider here O'Brien et al. studies of the influence of context and language on the status of understanding of the mathematical idea of implication as it is used in invalid inference schemes (AC, DA) by middle-class suburban school children in Missouri.

When instead of using "if p then q" language, the equivalent expression* "at least one of the following: -p, q" was used, subjects scored substantially and universally higher (O'Brien et al., 1971). In particular, growth from grade four to grade ten was very clear, yet even in grade ten only 50 to 60 percent of the responses to the various questions were correct (10 to 26 percent in fourth grade). Further subdivision of the questions, in both language forms, into items of which the content was defined as class inclusion in context, and items of which the content was defined as causal in context, resulted in substantial differences in subjects' performance. They tended to favor causal items in grades six, eight, and ten, whereas the fourth-grade subjects performed almost precisely the same on the two groups of items, with both overall and consistent misinterpretation of $p \rightarrow q$ as $p \leftrightarrow q$ at the 80 percent level. This effect occurred in all but MP logical forms. In cases when the logically equivalent language of "at least..." was used, virtually no differences were noted between contexts in either overall or consistent misinterpretation of the conditional sentence as biconditional.*

---

*See section 1.3.4, case (iii), page 15.

The effect of context, as investigated in the above study, which occurs in all logical forms but MP, suggests that high school subjects do not apply formal reasoning to the task of detecting the necessity of conclusions in inference patterns involving if-then statements.

1.4.7 Negation. Forms of negation were found by Suppes and Feldman (1971) to be much harder for students than other sentential connectives (namely: and, or). Hill, in her previously cited study, found similarly that negation, when added to the standard form of a principle of logic, increases the difficulty in making valid deductions utilizing this principle.

---

*For critique of the distinction between class inclusion and causal content of a conditional statement see section 3.1.1.

Wason and Johnson-Lair, in a book (1972) summarizing their comprehensive research, devote the first five chapters to negation. "Our experiments suggest," they say, "that negation does involve an extra step, or mental operation, and that when the negative lacks a preconception, such a step tends to be deliberately and consciously performed." This deliberate performance is indicated by the extra time needed for processing information given in a negative way. "It is as if the affirmative preconception has to be recovered before the meaning of the negative can be grasped.... On the other hand, in everyday life this extra step goes unnoticed because the preconception has already been processed as part of the context of the utterance." Other variables which were tentatively suggested by these psychologists as affecting the process of understanding were: (a) the possible emotional connotations of negative terms derived from their association with prohibitives which may, at least momentarily, inhibit response, and (b) "The scope of negations in terms of whether they are sentential or constituent may affect their grasp as a function of the specificity of their inference."

Apart from the acknowledged difficulty of understanding negatives, there is a special difficulty which arises when they occur in deductive arguments. A sequence of experiments reported by Wason and Johnson-Laird in their book points out that "as a general rule, there is no particular problem when they deny affirmative propositions - an explicit negative may, in fact, be easier than an implicit negative - but when a negative is itself denied by an affirmative, it becomes difficult to keep track of the argument. And the most parsimonious explanation for this seems to be the difficulty of 'double negation.'"

Wason and Johnson-Laird worked mostly with college students. Application of their findings to younger children is therefore of interest. Chen's (1975) findings go along these lines. In his cross-age study of children at grades five, eight, and eleven, he stated in the negative statements and the number of correct responses stated in the equivalent positive statement. Chen interpreted this as resulting from the fact that the negative statement affirmed the information given in the problem, whereas its positive equivalent required a transformation of this data.

Paulus (1967) studied children's ability not only to judge validity of deductions, but also to actively deduce. He reported that on both the assessing and the deducing forms of his test of conditional reasoning, the items containing negation were not more difficult than the others.

Roberge (1969) studied the same problem within concrete-familiar content only. His results indicated "that negation in the major premise had a marked influence on the devleopment of logical ability in children." This effect was consistent across fourth, sixth, eighth, and tenth grades. "Although negation had an influence on children's reasoning for both class and conditional reasoning, it apparently had a stronger influence in class reasoning." He suggested further investigation, which O'Brien tackled, of the effect of negation within specific principles of inferences.

O'Brien (1972) studied the effect of the negation mode of a conditional sentence, and its interaction with inference forms: MP, MT, AC, DA, controlling for item-content effect on form and negation mode. The study was conducted in a girls' high school in Saint Louis, grades nine through twelve. DA and AC were widely regarded by subjects as valid inference patterns. The relative number of correct responses for these forms was: MP > MT > DA > AC, in a range of 95.19 percent to 11.16 percent. Causal items were again found easier than class inclusion ones, with no difference among overall means for the latter and for random and nonsense content. The interaction of logical form with negation mode gave the following picture in number of correct responses:

For MP: ++ > -- > +- > -+, in a range of: 91%-98%
For MT: ++ > +- > -- > -+, in a range of: 51%-72%
For DA: +- > -- > -+ > ++, in a range of: 30%-37%
For AC: +- > -- > ++ > -+, in a range of: 8%-13%

The fact that -+ negation mode comes last in three out of the four forms gives support to Wason and Johnson-Laird's findings that "... abstract material, difficulty of the task, situations which are strictly binary..., the negation of the antecedent in $p \rightarrow q$, and conditional rules, are factors which are likely to lead to the conversion of conditionals," and hence to difficulties in deriving the right conclusions.

1.5 Methods Previously Employed in Teaching Logic and Those of the Present Study

The purpose of the present study was to promote young children's conditional reasoning through systematic teaching. There, fore, previous studies of attempts to approach young children with logical activities are of particular interest here. These are reviewed in this section and compared with the present study.

1.5.1 Previous attempts to teach logic or conditional reasoning.

(a) White, 1936.

Subjects: 2 classes of boys. Mean age 12 years and 11 months.

Instructors: The experimenter only.

Instruction: Three months. Once-a-week instruction in logic in addition to regular program in grammar.

Results: Class receiving the lessons in logic scored significantly higher on a reasoning test as well as in composition and English-construction test.

(b) Morgan and Carrington, 1944.

Subjects: Second through sixth graders. Control vs. experimental group at each level.

Instructors: Experimenter only.

Instruction: Experimental groups received graphic demonstrations of solutions to ten syllogisms given previously as an unexplained test.

Results: "Our results seem to show that the critical periods for learning relational syllogisms are the third and fourth grades, and that the graphic instruction is a factor in facilitating this learning."

(c) Hyram, 1957.

Subjects: "Two paired and equated groups of 33 children each." Mean age: 14.0 years, mean IQ: 106.18. Seventh and eighth graders.

Instructors: Experimenter in the experimental group.

Instruction: Four months of 250 minutes per week of instruction "in the development of the following seven concepts of logical thinking:

a. The Nature of Thinking in General
b. The Tools of Thinking
c. The Nature of Definition
d. The Nature of Eductive Inference
e. The Nature of Deductive Inference
f. The Nature of Experimentation
g. Common Errors in Reasoning."

This instruction was based on five hypotheses, three of which are quoted here:

"2. That logical thinking is no more than the application of the rules of logic to factual data in order to arrive at valid as well as true conclusions. It follows from this assumption that an individual's growth in the ability to do logical thinking must depend upon his acquiring a working knowledge of the basic rules of logic..."

"4. That the most effective way... is through direct instruction."

"5. That direct instruction should consist of a. Materials and learning content which embody the principles of logic. b. Teaching methods that provide full opportunity for the pupil to discover for himself these principles and to formulate them as generalizations."

Results: "It is highly feasible to conclude that the Experimental Group was superior to the control group in final reasoning ability as measured by the original test."

(d) Suppes, 1962.

Subjects: Experimental classes of selected fifth grade students. "Approximately 25-30 percent of the fifth graders were selected by the administrators on the basis of general ability and high achievement in mathematics."

Instructors: Classroom teachers from regular school staff.

Instruction: Using the book "Mathematical Logic for the Schools" by P. Suppes and S. A. Hill, pretrained teachers taught for a full year "at the pace they felt to be most appropriate for their classes." The text introduces symbolic representations in Chapter 1, rules of inference and truth tables are introduced later through examples and are presented in the symbolic formal way typical of mathematical logic.

Results: Evidence from the one pilot class reported to "indicate that its level of accomplishment was comparable to that of a college class in mathematical logic, although its pace was much slower."

(e) Suppes and Binford, 1965.

Subjects: Same as above, a year later, i.e., sixth grade, plus 12 new fifth grade classes.

Instructors: Same as above.

Instruction: Second-year students completed between 162-284 pages of the above text. First-year students completed 117-183 pages. Symbolic manipulations involved in derivations was a part of the program.

Results: "The upper quartile of elementary school students can achieve a significant conceptual and technical mastery of elementary mathematical logic. The level of mastery is 85 to 90 percent of that achieved by comparable university students."

(f) Ennis and Paulus, 1965.

Subjects: Groups of fifth, seventh, ninth, and eleventh graders.

Instructors: Researchers only.

Instruction: One period a day for fifteen days. Conditional logic was taught throughout this period.

Results: "...there is not much point in trying to teach conditional logic in elementary and lower secondary."

(g) Miller, 1968.

Subjects: A seventh grade class.

Instructor: Researcher only.

Instruction: "After completing the (pre) test, the students were interested in knowing the correct answers. In the discussion which followed the form of the patterns was investigated." There were 12 fifty-minute class periods in which "concepts were introduced in terms of physical world situations and then abstracted and symbolized. The laws of a two-valued logic were introduced as rules, much like... when one plays a game." The basic components were meaningful sentences to Wisconsin residents.

Results: "At the conclusion of the unit students were able to correctly test the validity or invalidity of an inference pattern."

(h) McAloon, 1969.

Subjects: 26 classes, 13 of third graders and 13 of sixth graders; all of average IQ or above.

Instructors: Classes were randomly assigned to 25 teachers for the four different treatments (instruction modes).

Instruction: Four different modes:

(i) Logic interwoven with mathematics, by teachers pretrained in logic.

(ii) Logic separated from mathematics, by teachers pretrained in logic.

(iii) No logic, by teachers pretrained in mathematics.

(iv) No logic, by teachers who received no in-service training.

Results: In both grade levels, groups (i) and (ii) scored higher on class and conditional reasoning, with no significant difference between (i) and (ii).

(i) C. Carroll, 1970.

Subjects: Selected ninth graders with low mathematics achievements, but above the 25th percentile in reading. 72 girls of whom 24 were taught logic, and 100 boys of whom 24 received instruction in logic.

Instructor: Researcher.

Instruction: Conditional reasoning was taught once a week for 39-42 minutes per period during 7 weeks. The method of teaching was oral discussion led by the instructor.

> "The general pattern followed in the presentation of each form of argument was to familiarize the student first with the basic forms through the use of arguments based on concrete objects or familiar subject matter. Then, gradually, attempts were made to have the students abstract the form from the content and judge the validity of the form on its own merit."

The forms of arguments were introduced one at a session in the following order, MP, AC, DA, MT (see section 1.3 for these abbreviations). The forms were imbedded in content defined as concrete familiar, symbolic, misleading, and removed from reality.

Results:

> "The improvement among subjects in the experimental groups measured by the percent of the subjects whose total score was higher on the posttest than on the pretest was not significantly greater than that among students in the control group."

Out of the four logical forms taught, only for AC was there a significant difference between the percent of subjects in the experimental groups who improved and that of subjects in the control group who did.

(j) Weeks, 1970.

Subjects: 30 second- and 30 third-graders; half of each were randomly assigned to the experimental group.

Instructor: Researcher only.

Instruction: 3 half-hour sessions a week for eight weeks. Dienes' suggestions for use of the attribute blocks were followed with slight modifications. "The primary role of the investigator was to introduce games, to encourage subjects, to stimulate discussions, and to help summarize discoveries." Topics covered (through the use of attribute block solely) were acquaintance with the blocks and forming of sets according to attributes (first nine meetings), "and" "or" "not" conjunctions (next five meetings), "if p then q" and its logical equivalent, "either not-p or q" (next four meetings), transformation, universal and existential quantifiers, valid conclusions from given premises (in the last six meetings).

Results: Attribute-block training had a strong, significant, positive affect on the development of logical reasoning ability and perceptual reasoning ability in both grade levels, with no significant difference between the two levels. These findings were based on the researcher-developed test in logical reasoning which consisted of 36 items "each in the form of two or three verbal premises and a possible conclusion presented as a question" similar to O'Brien's (1968) and Hill's (1961) tests. All items have content remote from attribute blocks, representing logical form of sentential and quantificational logic including four connectives common to mathematical logic ("and" "or" "not" and conditionals); some items require transitivity of the conditional, others call for MP, MT, AC, DA as well as classical syllogisms with one- and two-place predicates.

1.5.2 The present study

Subjects: Two fourth-grade and four fifth-grade classes of average and above-average general ability.

Instructors: Current classroom teachers of these classes.

Instruction: 20-25 regular class sessions, 4-5 times a week, using researcher's developed material on MP, MT, AC, DA in conditional reasoning embedded in realistic or hypothetical reasonable content. As far as language was concerned, statements were limited to simple ones where only negation and conditional connectives occurred. Manipulative aids in a game atmosphere were used to provide a variety of concrete situations in which the relevant logic appeared. No direct formulation or teaching of the rules of logic took place. Discovery of the underlying rules was left to the individual students who may have sensed the existence of the rules intuitively through the rich experience provided in many similar cases, but this discovery was never forced upon them. The use of symbolic representation was limited to abbreviations of particular names in discussion. This was intended to provide hints for abstraction of the logical forms. (Results - see Chapter 5, page 122.)

1.6 Teachers' and Prospective Teachers' Conditional Reasoning

Because employment of the experimental materials by ordinary teachers was one goal of this study, teachers' mastery of logic was of major concern. As Eisenberg and McGinty (1974) stated: "If elementary school teachers make the same types of errors in logical reasoning as elementary school children, then how can one expect to achieve in the schools the goal of critical thinking?" In their study comparing and contrasting the error patterns that prospective elementary school teachers and second and third grade students make in a logic test, they found that for the overall test the mean score of the prospective teachers was significantly higher than the mean score for the elementary school students; however, these mean scores were only 53.93% and 54.98% right answers for the first and second course for prospective elementary school teachers, respectively.

Moreover, when one focuses attention on the specific sentential logical forms of items within the test, troublesome areas become evident. On AC items, scores of the college students did not differ significantly from chance. For items of the form "P or Q; P, therefore not necessarily Q," and "P or Q; not Q, therefore P," which can be interpreted as DA and MT items, respectively, if we change "P or Q" to the logically equivalent statement "if not P, then Q," the mean scores for preteachers and for children were not significantly different. Indeed, for the first kind all scores were significantly lower than a chance score.

R. McCoy (1971) conducted a study of the effects of three different strategies of proof instruction on college students majoring in elementary education. 139 students participated for four weeks in his main study. Four instructional plans were used: (i) formal logic; (ii) applying the rules of logic to proof construction; (iii) applying the rules of logic, by model building, to proof construction; and for control, (iv) no instruction on proof construction. All the strategies of proof instruction were taught by the author with the exception of the control group. A lecture-response approach to the teaching was used.

He found the post-unit mean scores on proof construction to be (i) 22.1 (ii) 27.8, (iii) 27.3, (iv) 18.4, out of a maximum 36 with respect to the above methods. These results support Suppes and Binford's (1965) conclusions that: "The more dedicated and able elementary school teachers can be adequately trained in five or six semester hours to teach classes in elementary mathematical logic."

The need for training teachers prior to their employment of logic in their instruction was recognized by Suppes (1962, 1965) and McAloon (1969) who, as mentioned previously, conducted experimental teaching in logic through regular elementary school teachers.

The findings cited above indicate an urgent need for teachers' training if effective handling of materials by teachers is sought. This need also brings this whole psychological background chapter to a close at the point where it opened, namely, to use Eisenberg and McGinty's words: "The conjecture cited by Donaldson (1963), Gardiner (1965), Hill (1961), and others that maturation comprehensively affects sententially logical thinking is not supported.... Remediation in these areas will not come without work. Maturation is not enough."

1.7 Questions Studied

1.7.1 Objectives. The preceding disucssions in this chapter have focused on previous treatments of principal problems and on background for the particular objectives of the present study. The following is a list of these objectives:

a. To develop a unit in conditional reasoning for the upper elementary grades, aimed at familiarizing students with conditional reasoning paradigms without explicitly teaching either the rules of formal logic or any algorithmic solution. This unit should be based upon formal logic as discussed in section 1.3, and should utilize the experience of previous attempts as discussed in section 1.5. Conditional logical forms in the developed unit should be limited to content expressed in simple sentences, where the only connectives included are the negation and the conditional ones. Conditional reasoning in this unit should be limited to familiar content in concrete or hypothetical situations that do not contradict everyday experience and do not have negative emotional connotation as discussed in section 1.4.

b. To develop this unit to the point where ordinary teachers will be able to apply it in regular classroom settings, despite the acknowledged difficulties as discussed in section 1.6.

c. To develop this unit to the point where, when applied by regular teachers in their ordinary classes, a marked progress in students' conditional reasoning ability will be obtained.

d. To develop a reliable instrument of measurement to determine change in students' conditional reasoning ability.

1.7.2 Hypothesis tested. In general, it was hypothesized that the objectives stated in section 1.7.1 are attainable. Namely:

a. There exists a proper method of introducing fourth and fifth grade students to conditional reasoning on an intuitive level, with the limitations stated in part a of the previous section.

b. Some elementary school teachers are able to provide instruction in conditional reasoning in their regular classes after receiving a proper training.

c. Students of different fourth or fifth grade classes in any one school district that does not have a tracking system will exhibit similar performance on a pretest level. This performance will be high on MP and MT logical forms and poor on AC and DA logical forms.

d. Negation in the first premise, i.e., in the conditional sentence, will increase the difficulty of any of the four logical forms.

e. Classes introduced to conditional reasoning by their teachers will exhibit in a posttest a marked progress in AC and DA logical forms. The level of observed performance on MP and MT will remain equal or slightly higher than that of the pretest.

f. Performance of students in classes not introduced to conditional reasoning will remain unchanged, if the time lapse between the pretest and the posttest is no longer than a few months.

g. Students' errors in both pretest and posttest will conform to the model of error prediction described in section 1.7.3.

1.7.3 Model for error prediction in conditional reasoning. After considering the explanation of errors by atmosphere effect, conversion effect, probabilistic reasoning, and considering "if... then" implying its converse or as "if and only if," the present study assumed and tested the model in table 5.1 for error prediction due to fallacious conditional reasoning.

Another way to interpret the pattern of mistakes given in table 5.1 was suggested by L. Henkin (1974). He conjectured that the tendency to wrongly affirm the antecedent in the fallacies of the form Affirming the Consequent, and to wrongly deny the consequent in the fallacies of the form Denying the Antecedent, has nothing to do with logical or illogical thinking,

Table 1.5 Expected Errors in Conditional Reasoning

	AC	Corresponds to	DA	Corresponds to
++	clues: $p \to q$ q question: p? expected error: YES	MP in ++ or MT in --	clues: $p \to q$ $-p$ question: q? expected error: NO	MT in ++ or MP in --
+-	clues: $p \to -q$ $-q$ question: p? expected error: YES	MP in -+ or MT in -+	clues: $p \to -q$ $-p$ question: q? expected error: YES	MT in -+ or MP in -+
-+	clues: $-p \to q$ q question: p? expected error: NO	MP in +- or MT in +-	clues: $-p \to q$ p question: q? expected error: NO	MT in +- or MP in +-
--	clues: $-p \to -q$ $-q$ question: p? expected error: NO	MP in -- or MT in ++	clues: $-p \to -q$ p question: q? expected error: YES	MT in -- or MP in ++

Note: The right answer in all cases above is: Not Enough Clues (NEC). The above error patterns coincide with the logical analysis of errors given in the last paragraph of section 1.3.6, in which MP is applied to the converse of the first clue in the AC cases, and MT is applied to the converse (or MP to the inverse) in the DA cases.

but may result from some sense of "language balance." If he is right, then a major part of the right answers previously observed by researchers in MP and MT cases, may not indicate logical thinking after all. Instead, the same almost impulsive sense of language balance may lead to rightly affirming the consequent in MP cases, and to rightly denying the antecedent in MT cases. Whether this explanation has roots in actual experience could be verified through an examination of arguments and explanations subjects give to justify their answers, and by the extent to which their consistency in giving right answers to MP and MT items is persistent even after becoming aware of the fallacies. Another way to test this explanation is by changing the connective to see whether this sense of "language balance" leads to wrong answers in cases similar to MP and MT.

Predicted errors and the various explanations for them provided the basis for development of the experimental teaching materials. It became evident that efforts should be made to demonstrate what is wrong with fallacious reasoning through counter examples and careful argumentation.

CHAPTER 2

THE EXPERIMENTAL UNIT

Chapter Overview

An overview of the experimental unit in its final form is given in section 2.1. This section includes an introduction (2.1.1), in which the approach is analyzed and compared with previous studies. The teachers manual for the unit, including students completed paper-pencil worksheets, can be found in appendix 7.1. The rest of the chapter describes the development of the experimental unit up to the final version used in the main study. The main study itself will be described in Chapter 4 (page 112).

The course of development involved several cycles of teaching-revision-reteaching. The teaching trials consisted of three major phases: (i) individual work with two fourth graders (section 2.2), (ii) small-group work with fourth graders, fifth graders and with one mixed age group (section 2.3), and (iii) whole-class trials with two fourth and one fifth grade class (section 2.4). The experimenter conducted the first two phases. The third phase, which constituted the pilot study, was conducted by regular classroom teachers. Lessons learned from each phase, modifications of teaching strategy, revisions of materials, and further development are described after each trial.

2.1 The Experimental Unit - An Overview

2.1.1 Principles of the approach. Because the present study was designed to provide an introduction to conditional reasoning for upper-elementary-grade students, this researcher decided to consider only familiar content. It would often be concrete in nature but could sometimes be of a hypothetical nature. Still, it would never be contradictory to facts or to previous experience. It was also considered desirable to keep the content neutral from emotional reaction and to limit symbolic representations to abbreviations of terms in current discussion.*

Many of the studies reviewed in section 1.5 exposed subjects explicitly to formal logic in symbolic representation. The purpose of the present study was to teach a proper use of "if...then" in any meaningful context. Achieving this purpose using symbolic representation requires transfer from the formal language to terms of specific content, which cannot be expected from elementary school students. Also, symbolic representations appear inadequate for elementary school in the light of Piagetian theory. If generalization is sought, Hyram's (1957) approach of inductive discovery appears to be the most appropriate at all levels because of the educational value of transfer and retention. However it should be remembered that fourth and fifth graders may not be able to express in a symbolic way such discoveries of general logical patterns. Moreover, the value of symbolic presentation to them is questionable.

Despite the previously mentioned negative perspective for young children's ability to learn conditional reasoning, indicated by Ennis and Paulus (1967), Weeks reports a significant success in working with very young children. The use of attribute blocks by Weeks (1970) as the only manipulative aid raises questions of maintenance of interest as well as of transfer. As Ellis (1965) points out, transfer is facilitated through the use of a variety of tasks in the learning situation. Moreover, MacGinitie and Ball (1968) emphasize that an essential condition for transfer to occur is for students to recognize the common features underlying the variety of tasks. Students' self-discovery of the common features leading to possible generalizations seems to Brownell (1936) to be a necessary condition for transfer. Weeks' verbal-testing results after the twelve hours of attribute-block training of second and third graders are surprising not only in the wide range of logical skills achieved, but also in view of the transfer required. "We can expect more transfer when the training task and the criterion task resemble each other in their overall characteristics" - Stephens (1963) reports, but in Weeks' case the content is so different that the transfer of logic from the training task to the test tasks is worth pointing out. In the present study a variety of manipulative aids were used to provide a rich, concrete experience as a basis for execution of deductive thought in a meaningful environment, and to increase the likelihood of transfer.

---

*For further discussion of practical problems related to content, see section 3.1.1, page 84.

In several high school text books there are sections devoted to logic, but because these are intended for students older than those involved in the present study they will be reviewed only briefly for comparison with the approach taken here.

Some high school mathematics projects (e.g., Exner et al., CSMP, 1972) start their curriculum with direct teaching of logic.

It was not the purpose of the present study to teach formal logic, nor to analyze the nature of the deductive process itself. Rather the purpose here was to teach the process of making valid deductions by providing experience in making valid deductions and practice in avoiding nonvalid deductions as well. The decision was made to postpone the formal analysis until after students were exposed to a variety of concrete examples and executed logic on an intuitive level.

There are some new textbooks that devote a separate chapter to teaching the process of constructing a proof. R. Davis (Madison project, 1964) leads junior high school students to shortening a given list of sentences, in other words to draw a conclusion from this list. This seems to be a very worthwhile idea, even though the various lists he gives do not lead to any generalizations or discrimination between valid inferential techniques and non-valid ones.

Another well known approach to acquainting students with ideas of mathematical logic is to teach the construction of truth tables for sentences built with propositional connectives (e.g., Suppes and Hill, 1964). This task is essentially computational and algorithmic. The approach in the present study is different in that it is directed to leading students toward the intuitive construction of a deductive argument. Truth tables may be incorporated, in such an approach, as a technique for verification of the logical validity of a conclusion, which in turn will indicate the provability of the conclusion. But truth tables alone cannot serve the purpose of putting the students into a thinking process of distinguishing valid from nonvalid inferences; a truth table is an automatic algorithmic process that does not take any thinking. Also, truth tables have operational limitations. For example, it is impossible to construct a truth table for general sentences of the form, "For all x, x is...," because there may be models in which such a sentence is true and others in which it is false (see section 1.3.7). Teaching the technique of constructing truth tables and of finding such models, although it may attract young children, is time-consuming, and was not used in the experimental unit.

In all but two of the studies cited above in section 1.5, the researcher carried out the instruction. This limits the generalizability of these studies insofar as applicability to regular elementary-school instruction is concerned. Teachers usually do not possess an extensive background in logic or a rich experience in mathematical thought. Hyram (1957), cited previously, suggests that teachers should take basic courses in critical thinking at teacher-training institutes in an effort toward "bridging the gap between educational theoretical aims and the actual results of teaching practices." In the present study, actual classroom teachers took the active role of teaching. These teachers participated in a simple training program prior to their own teaching periods.

In the two cases where ordinary classroom teachers were the instructors (Suppes, 1962, 1965, and McAloon, 1969) there was departure from the regular classroom setting at school. In one case classes were formed on a selective basis, and in the other teachers were randomly assigned to classes. Even though in each case there were probably good reasons for the procedures adopted, they again limit the generalizability of these studies to regular elementary school situations. In the present study, classroom teachers taught their regular classes in an ordinary school setting.

In some of the above studies the treatment seems to be somewhat short and isolated. This may explain at least in part Carroll's (1970) inability to show significant differences between the treatment and control groups despite her very reasonable teaching methods. In the present study teaching occurred four to five times a week and lasted for five to six weeks.*

The teacher edition of the experimental unit is attached in appendix 7.1. It includes answer sheets for those parts of student-activities involving worksheets. The rest of this section conveys the general structure and content of the experimental unit..

The unit consists of seven chapters: Electric Cards; Dominoe Activities; Pictorial Activity; Numbers and Their Properties; Playing Cards; Colored Light Switch Box; Prepare a Quiz. Each chapter is a set of small-group activities introduced by a teacher/whole-class activity. The first set, Electric Cards, is a motivational activity for the whole unit and it also leads to the final project.

Between the first introduction and the final project, the unit is subdivided into three parts according to the objectives to which each set of activities addresses itself. The first part is designed to demonstrate the implications of a conditional sentence; the second part shows how conditional sentences are derived from other sentences; and the third part integrates the first two in order to produce some generalizations.

---

*See also section 1.5.2, page 50.

2.1.2 Motivation and final project. The first set, Electric Cards, sets the motivation for learning the whole unit through a self-conducted group activity of problem solving. There are about 300 problems; each is typed on a separate card. An immediate checking and feedback system is available through an electric tester. The tester lights a bulb when ond only when its terminals are attached to the metal contacts connected to the right answer.* The children are instructed not to test an answer before each group member has independently offered an answer and a discussion has taken place in case of disagreement. The final project at the end of the unit involves the whole process of producing students' self-made electric cards. The student starts by inventing a true conditional sentence, the converse of which is false. Next, the student makes up a problem for each card, wires the card and challenges a frient with it. (The term converse was not introduced; instead, students learned about a conditional sentence and its "flipped-over" mate.)

2.1.3 The first part. The Domino Activities and the Pictorial Activity** constitute the first part of the unit. These activities demonstrate by a variety of examples the idea that whenever a conditional sentence (if p, then q) is true or assumed, it excludes the possibility of p and not-q occurring in conjunction, and this possibility is the only one excluded by that hypothesis. In other words, the conditional sentence allows either one or more of the following to occur: p and q; not-p and q; not-p and not-q. In particular, the possibility in the middle is not excluded.* In the second part (section 2.1.4), the converses of these two theorems are demonstrated.

---

*See details in appendix 7.1, page 222.

**See appendix 7.1, pages 228, 236.

The Pictoroal Activity is a set of pictures and a conditional sentence. The task is to select those pictures that contradict the sentence.**

The conditional sentences and the pictures become more complex as students progress. In principle, the pictures given for a conditional sentence p → q illustrate four situations: p and q; p and not-q; not-p and q; not-p and not-q. (See diagram 2.1 for a sample.) Students are expected to discover that the pictures contradicting the conditional sentence are those where p and not-q are illustrated. The understanding that this is the only case that contradicts the sentence is developed through group and class discussions.

For the Domino Activities, the teacher makes up a story about Paul who keeps losing his dominoes. One day Paul discovers that the dominoes he still has obey a rule of the form: If there is a 3 on any of my dominoes, then there is a 2 on it. Students study some incomplete domino sets to find out true rules for them, then

---

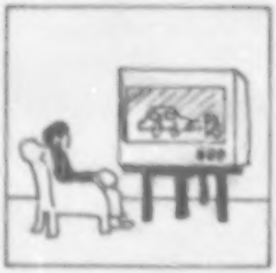
*The logician would recognize the two theorems behind it: (i) $p \to q \models -(p \wedge -q)$; (ii) $p \to q \models (p \wedge q) \vee (-p \wedge q) \vee (-p \wedge -q)$.

**For a discussion of why finding the contradictory picture is easier than finding the ones that agree with the sentence, see the teachers guide for that activity (appendix 7.1, page 240).

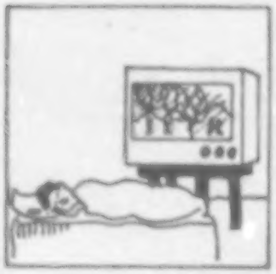
Diagram 2.1 A Sample of the Pictorial Activity

In each picture you see a man doing something, while the TV is either on or off. Read the conditional sentence at the bottom of this page and write the number of each picture which disagrees with it.

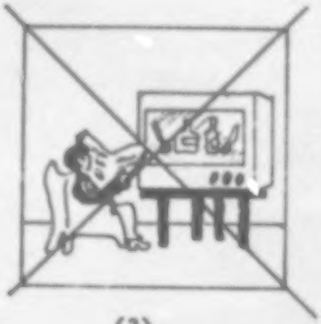
3



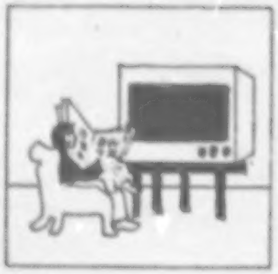
(1)



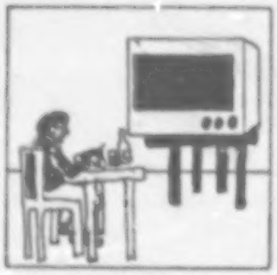
(2)



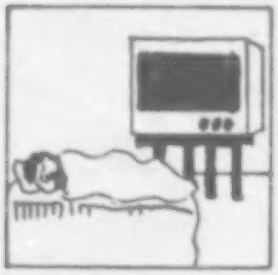
(3)



(4)



(5)



(6)

If the TV is on, then the man is not reading.

This is the first conditional sentence, with negation, in this activity. Student will as usual be called to describe the pictures in terms of: TV is or is not on; the man is or is not reading. Concentrate on those pictures where TV is on and find the picture which does not agree with the sentence which is the one where the man is reading while TV is on. Pictures 4,5,6 do not disagree with the sentence!

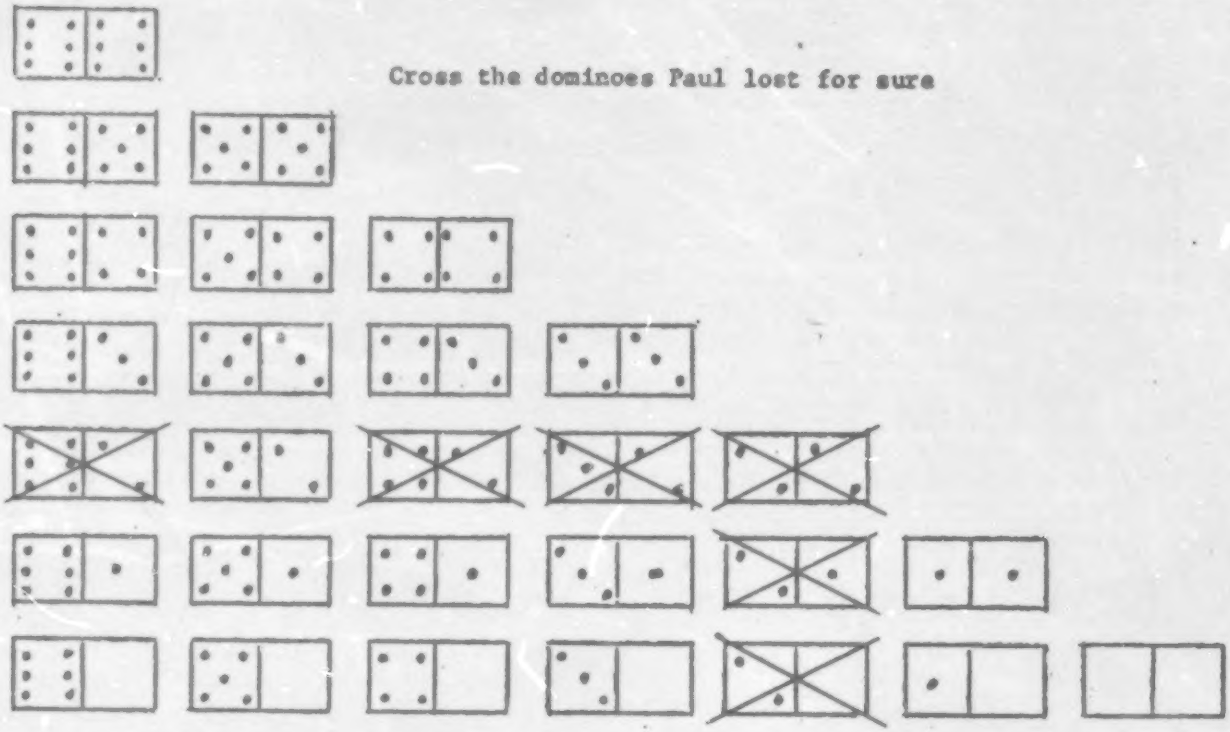
they play "Paul, the scatterbrained guy" where students take turns being Paul and the one playing Paul puts an incomplete domino set on the table. His peers have to guess his rule. (There may be more than one true conditional sentence for a given incomplete domino set!). Later on students cross out on a chart of a complete domino set those dominoes Paul definitely lost for a given rule of an incomplete set. E.g.: if there is a 2 on one side, then there is a 5 on the other side (2 → 5, in short). For this rule a student should cross out all the dominoes that show 2 but do not show 5. Notice that only these dominoes must be missing from Paul's set. He may or may not have lost a dominoe like 5:3. On the other hand students do the same thing for the "flipped over" mate of the above conditional sentence, in other words they cross off the chart (a clean copy for every rule) those dominoes Paul definitely lost if the rule for his incomplete set is 5 → 2. A comparison of the answers for the two problems demonstrates the difference between a conditional sentence and its converse (see diagram 2.2, page 68).

The dominoes are used also for another purpose; to bring about the realization that a conditional sentence carries information which is identical to that carried by its contrapositive mate (i.e., its inverse).

2.1.4 The second part. The second part of the unit -- Numbers and Their Properties, and Playing Cards -- serves to illustrate to the student, again through concrete examples, the converse of each of the theorems discussed in the first part, namely: (i)

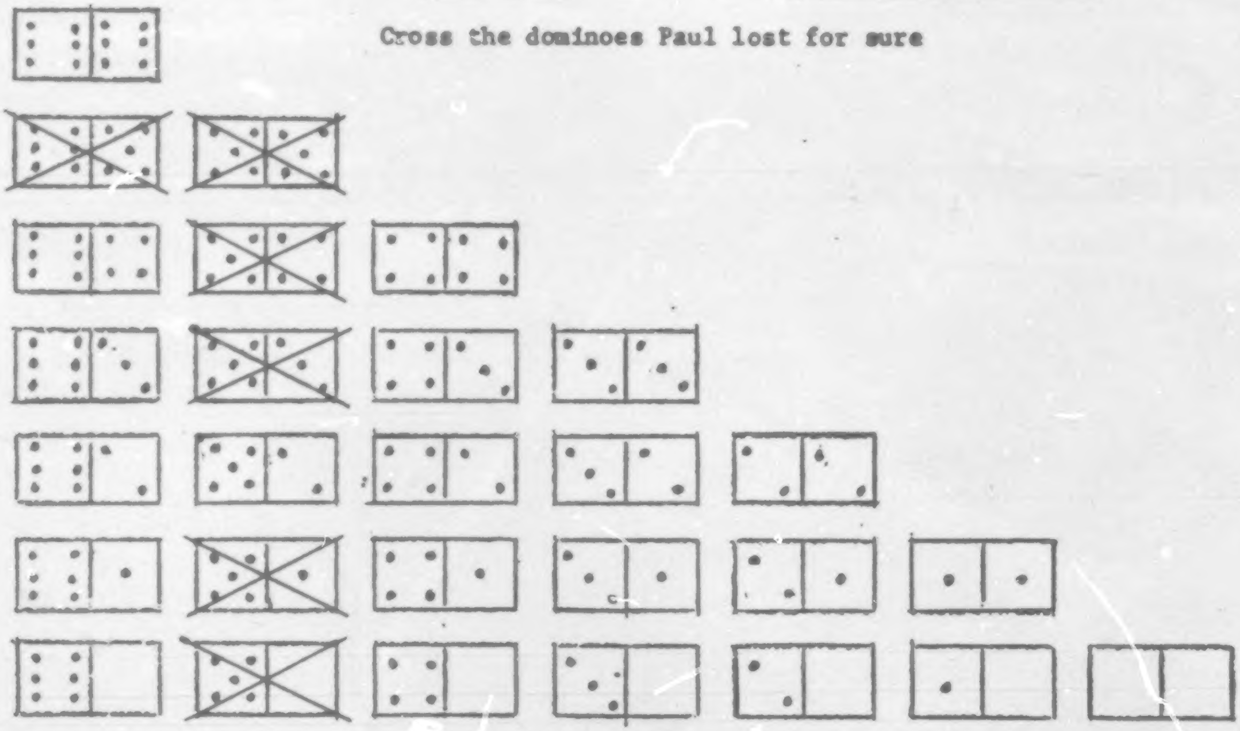
Diagram 2.2 A Sample of the Domino Activity
(Illustrating the independence of two converse sentences.)

Paul's rule for his incomplete dominoe-set is 2 → 5



Cross the dominoes Paul lost for sure

Paul's rule for his incomplete domino-set is 5 → 2



Cross the dominoes Paul lost for sure

$-(p \wedge -q) \models p \rightarrow q$ (ii) $(p \wedge q) \vee (-p \wedge q) \vee (-p \wedge -q) \models p \rightarrow q$. Combined together, the first two parts give concrete experiences which demonstrate the equivalence of a conditional sentence to either the disjunctive or conjunctive statements above. To take away any doubts it should be noted again that the abstract language used above is not the language used in the classrooms.

In addition, the second part of the unit, particularly the activities with playing cards, provides a rich source of examples for the equivalence of a conditional sentence and its contrapositive. This recognition of equivalence and of the logical independence of two converse sentences was regarded as prerequisites for development of recognition and understanding of the validity and invalidity of inferences made from conditional premises.*

In Numbers and Their Properties the work is mostly on an individual worksheet basis. Students are to position numbers in a 2 × 2 matrix under the proper categories of p and q, p and not-q, not-p and q, or not-p and not-q, again of course with particular p and q, e.g., p is "the number is less than 30," q is "the number is less than 60." For this example students would have a hard time finding a number for the box designated by p and not-q because there is no number less than 30 and not less than 60. Part of the task is to discover this fact and then rephrase it as a conditional sentence: if a number is less than 30, then it is less than 60. Relations other than -- less than -- are suggested to the teacher in the manual (see appendix 7.1, page 286). Note: in the contrapositive: If a number is not less than 60, then it is not less than 30, students (and teachers!) tend to say "greater than" instead of "not less than", ignoring the equality case.

---

*See section 1.3 for underlying logic.

For the Playing-Card activities each group of students receives a deck of regular cards and a chart like the one in diagram 2.3. Students take turns in putting cards in place, explaining each time why they do so, e.g., for 7 of spades they are supposed to say: it is not red and not a heart, so it belongs here (bottom right box). Eventually, usually before all the cards are distributed, they realize that there would be an ampty box, in this case the upper right one. This box will stay empty because there is no heart card that is not-red. Rephrasing it in two ways as a conditional sentence calls again for using the contrapositive. It is particularly easy to see that the same chart can be described in two contrapositive ways if one starts the conditional sentence

Diagram 2.3 A Sample Chart for a Playing Card Activity

Put each card in the right place. What do you discover?

	Red	Not-red
♡		
not-♡		

Phrase your discovery as a conditional sentence in two ways.

once from the left marginal titles: "If a card shows a heart, then it is red," and once from the upper marginal row: "If a card is not red, then it is not a heart."

In the majority of cases up to this point, whenever a conditional sentence is discussed, four types of logical problems are formulated. The students answer these problems and provide their reasons based on their observations. At many points students are advised to take the teaching role and lead the class in a discussion of the analysis of some puzzles.

2.1.5 The third part. The third and last part of the unit, which includes activities with a colored-light switch box, more paper and pencil work, and preparations for the final project of the Electric Cards production, attempts to lead the students towards generalizations through sorting of the problems they dealt with previously and analyzing similarities and differences between the four kinds of problems.

Throughout the second and the third part of the unit, many abbreviations are used. Students like them because they save writing. What they may not know is that these abbreviations also emphasize the syntax of the various logical forms. They stand for constants in the logical sense, and thus serve as an intermediate step before variables take their place for the general pattern of the four relevant logical forms.* There is nowhere a direct teaching of the algorithmic solution discussed in section 1.3.8 (see page 23). However, sense of its existence is expected to emerge from these activities, even though it is never forced. The remainder of the chapter gives a detailed description of the writing-experimenting-rewriting cycle that produced this final version.

---

*See table 1.2, page 20, for a reminder of the logical forms.

2.2 Unit Development - Experimenter's Work with Individual Students

2.2.1 Exploratory work. Because several studies suggest that fourth grade is a promising starting point for a systematic introduction of certain patterns of logic (see section 1.4.2, page 27), the very first attempt to devise teaching strategies was made through individual sessions with two fourth grade students, a girl aged nine years and eleven months and a boy aged nine years and six months. These two children placed in the top quartiles of their classes, but were not the typical students; both were foreigners, had been staying in the U.S. for eighteen months at that time (April 1974), and English was not their native tongue (nor was it the experimenter's, by the way). Convenience due to personal relationships with the children was the main reason for choosing them for this part of the study.

Each child spent two or three 30-45 minute sessions with the experimenter. During these sessions the experimenter presented puzzles to the child. The puzzles were relevant items, of MP and MT logical form, from Hill's test (1961) and their modifications to undecidable items, of AC and DA logical form, as suggested by O'Brien (1968).* After a child gave an answer, his/her arguments for it were requested. When a wrong answer or a false argument were given, the experimenter conducted a question-and-answer dialogue to lead the child to a recognition of the mistake, and consequently to a change of the answer. Since the notion of -- not enough clues -- was not introduced in advance, expressions like: We can't tell, could be either way, not necessarily so -- were acceptable for undecidable items.

---

*For a review of these references, see sections 1.1 and 1.4.

The first session with each child was planned to last thirty minutes. After one-half hour, both children insisted on continuing. After forty-five minutes the experimenter still could not stop before promising to continue the puzzles again the next day. Both children seemed to become aware of certain differences among the puzzles which at first looked so much alike. They consistently answered correctly most of the MP and MT items and repeatedly committed fallacies in most of the AC and DA items even though each time they did, they later seemed to understand their mistake through the question-answer dialogue based on their incorrect answers or arguments.

In the second session, no real progress was yet apparent. Again it was easy, easier than before, to make the children realize what was wrong with their answers. "Oh, yes, it's the same thing, it does not have to be so and so, because it may also be so and so." This was a common pattern of reaction. Also, "Why am I so dumb?!? Let me try the next puzzle. Now, I'm sure I got it," but the next undecidable item was again answered spontaneously and very often incorrectly. The only difference was that as time passed, the very first clue, sometimes even given nonverbally by the experimenter, sufficed to bring the child to reconsider his/her answer and to come up with the right one with a well phrased, even if formally incomplete, argument.

The boy became tired, possibly frustrated after thirty minutes. He was no longer looking forward to a third session. So his next session was to take place when he asked for it, but he never did. However, often in his bed before he went to sleep, he generated new items, mostly of AC and DA logical forms, answered them correctly and gave surprisingly good reasons for their undecidability. Many of his ideas were later used in developing the Electric Cards activity.

The girl maintained her interest in the second session. She became bored in the third session, and there was a notable change in her performance. She had no more difficulty in getting the right answer to almost any puzzle of each of the four relevant logical forms. In the few cases she gave the wrong answers, she realized it right away, was able to explain her corrected answer, and volunteered an apology for the cause of her mistake. These apologies usually went along the lines logicians and psychologists discuss in their writings, namely, basing the definite answer on the converse of the given conditional sentence.* Experimenter's efforts to play the devil's advocate were rarely successful with this girl in the third session.

2.2.2 Basic lessons learned from the exploratory work.

(i) Providing experience with a rich variety of undecidable puzzles alongside decidable ones may prove an effective teaching strategy.

---

*See discussion of errors in section 1.3.6, 1.4, and 1.7.3.

(ii) It is necessary to find ways other than oral discussions to obtain and maintain students' interest for at least ten sessions of group work.

(iii) It is necessary to find ways to teach the difference between a conditional sentence and its converse.

(iv) Some fourth graders are able to learn how to judge a conclusion based on simple conditional logic inferences, and moreover they are able to express, in a clear way, their reasons for their judgment.

2.3 Unit Development - Experimenter's Work with Small Groups

2.3.1 Albany, Spring 1974. During May 1974 a group of three boys and four girls, all fourth graders, met for four weeks with the experimenter, twice a week for an hour. In the first and last meetings with this group an early version of experimenter-developed test* was used as both pretest and posttest; both tests were presented as boys vs. girls contests. The teaching program as a whole was not planned in advance, but rather on a session-to-session basis, using the experience learned in one session to plan for the next one. This allowed for high flexibility in exploring various teaching techniques and group work organization.

Attribute blocks (Elementary Science Study) and the colored-light switch box mentioned earlier were used as manipulative aids. Negation was gradually introduced into the examples.

In principle the order of teaching was based on introducing the logical forms one at a time, starting from MP, then adding AC (first two sessions), then DA (third session) and MT mixed with DA

---

*See section 3.1.2, page 90.

(fourth session), and finally integrating all four forms in the last two sessions.

The names MP, MT, AC, DA were not mentioned, nor were any syntactical considerations except those discovered by the students. Students never discovered the algorithmic solution, i.e., the relation between syntactical structure and right answer for an item as mentioned in section 1.3.8.

2.3.2 Basic lessons learned from this experiment.

(i) The method of introducing the logical forms one at a time in the order mentioned above was found less satisfactory in view of the high interest provoked in the last two sessions when all four forms were integrated.

(ii) This group of students seemed to enjoy most competitive games where one team challenged the other with its puzzles. This became a final part of every session in which at least two different logical forms were discussed.

(iii) The use of attribute blocks was found unsatisfactory because of prior use by the students and the remoteness from reality.

(iv) All seven students made considerable progress from the pretest to the posttest (from 41.5% to 57.5% right answers). Several items in the test were found misleading and later were changed (see section 3.1.2, 3.1.5).

(v) Students were confused by MT items. This led to a search for activities to introduce the distinction between a conditional sentence and its contrapositive. Consequently parts of Playing Cards activities and the Numbers and Their Properties activities were developed.

2.3.3 Berkeley, Summer 1974. During two weeks in July 1974 the experimenter met, for one hour daily, with ten students going into the fifth grade the following September. This was part of their summer school program at Berkeley, California. Students again were pretested and posttested. Right from the start, all four forms were introduced through examples with no explicit distinction among logical forms. It was hoped that some students would eventually discover syntactical differences and common features among certain problems. In fact four students did, and one of them even sensed the algorithm, even though he was unable to verbalize it properly. He got 31 out of 32 right answers on the posttest, and his reasons were mostly syntactical, for example he once said: "It's the other way around here, see the second line? It's like the end up there" (meaning the second clue is the same as the consequent of the conditional sentence) "so it (the answer) can't be yes and it can't be no, because..." and here he started to relate to the content.

The lessons learned from the previous experiment were incorporated into the teaching in this experiment. Materials developed between the two experimental teaching cycles (see section 2.3.2) were tried out and new ideas were brought up for modifying, extending, and replacing the old ones. Once again a flexible planning of the sessions allowed a great deal of experimentation, and the small size of the group allowed for close observation of these children's reasoning processes.

The basic mistake students made seemed still to result from their inability to separate the converse from the given conditional sentence and, as a result, they assumed that given $p \rightarrow q$ it was impossible to get not-p and q in conjunction.

2.3.4 Basic lessons learned and further development.

(i) Because students reacted negatively to paper and pencil work, this kind of work was reduced and whenever unconnected to activity, postponed to later parts of the unit.

(ii) A need for an attractive way of demonstrating that $p \rightarrow q$ excludes -- p and not-q -- and only it, was recognized. As a result Pictorial Activity was developed.

(iii) Students reacted favorably to manipulatives. A need for more activities with a game atmosphere was recognized. At the same time it was still desirable to enrich the unit with more examples of realistic situations where the four logical forms could be carefully used. Consequently the Electric Cards activities were developed.

(iv) Dominoes activities were developed to answer the need for demonstrations of the independence of a sentence and its converse.

(v) In late August 1974 after the additional activities were developed, a first draft for the teachers manual was written based upon the experimenter's notes during the Albany and Berkeley trials.

2.3.5 Fall 1974, Lawrence Hall of Science. Once again the experimenter worked with a small group of children, a mixed age group (7-11) this time. They met once a week for an hour and a half for eight weeks. This was an afternoon course, open to the public, given at the Lawrence Hall of Science at U.C. Berkeley.

The purpose was to provide demonstrations of the use of the experimental unit to teachers of the pilot-study which occurred simultaneously at Albany, California (see section 2.4). Unfortunately the teachers did not use this opportunity. It should be noted, however, that only half a session each time was devoted to the experimental unit and the other half to some other popular games. After all, students at this age group would not be expected to maintain interest in one subject for an hour and a half. The one-week time lapse between the two successive meetings was a great disadvantage to continuous improvement.

Most of the students were of high potential so that despite the long and discrete nature of the meetings and despite age differences their posttest scores were the highest obtained in either the pilot study or the main study.

2.4 The Pilot Study - Teachers Conducted Teaching

2.4.1 The sample and how it was secured. In August 1974 the superintendent of schools in Albany Unified School District, California, granted permission to conduct the pilot study in this district provided that teacher cooperation would be voluntary. He also agreed to grant three quarter-unit in-service district credit to each participating teacher who ran an experimental class.

Three weeks after the school year had resumed, two principals called a special fourth grade teachers district meeting. Five teachers attended; two agreed to participate in the pilot study. The other three agreed to let the experimenter pre- and post-test their classes for control, so that all the fourth grade teachers in the district participated in the pilot study either as experimental-group teachers or as control-group teachers. Among the teachers, three were males (one experimental and two control-class teachers), and two were female, (one experimental and one control-class teachers). The two experimental-classes were in different schools. For one of them a control class existed in the same school. The other two control classes were in a third school. All three schools are situated about a mile apart.

Albany is a small community in the San Francisco - East Bay Area. It is populated mainly by middle-class families and families of graduate students of the University of California, Berkeley. There were 49 students in the two classes of the experimental group and 75 students in the three classes of the control group. Procedures and results of the pretests and posttests are given in tables 3.1 and 3.2 of Chapter 3. The remainder of this section details the historical account of the pilot study, the problems encountered, and lessons derived.

2.4.2 Teachers' training. In the present study teacher training was consistent with the following recommendations of Suppes and Binford's (1965) that:

> "It is probably essential that this teacher training program be very closely geared to the actual program of instruction the teacher will follow in the classroom."

The experimenter spent one session of one and a half hours after-school time and five 30-35 minute sessions during lunch time with the two fourth grade experimental class teachers. Teachers were reluctant to meet after school and on weekends. Other school programs like swimming and field trips prevented frequent lunch time meetings. So the training period lasted three weeks. The teachers were presented briefly with each activity: Electric Cards, Dominoes, Pictorial Activity, Playing Cards, Numbers and Their Properties, Colored Light Switch Board, and Prepare a Quiz. In the third meeting the experimenter realized that the teachers were becoming less interested in the preparation and were eager to take the risk and start teaching on a daily planning basis. Therefore, rather than wait for the teachers to discover the algorithmic solution (section 1.3.8) it was shown to them but with instruction that the students should be allowed to come up with it by themselves. If the students do not, the algorithm should not be given to them. It was not a part of the teaching they were expected to do. The need to obtain the students' arguments along with each answer was also repeatedly clearly stated.

At the fourth meeting a copy of the first draft of the teachers manual was handed out to each of them and they were asked to do some reading. But they did very little. As a result, the teaching period started with insufficient mastery of the subject matter by the teachers and their insufficient acquaintance with the experimental unit as well. Starting from the second week of teaching the experimenter prepared a weekly plan in a written form. Each week's plan included details for each of the four periods expected to take place in that week with reference to the teacher's manual. This plan was partly previewed every Friday in the half-hour staff meeting (that took place during lunch time).

2.4.3 The teaching period. The teaching period started with a pretest on October 21, 1974, and ended in a posttest on November

25, 1974, altogether 5 weeks, each of four 30-40 minute class periods. (For the tests see section 3.1.) The schedule of classes was planned such that the experimenter was present in each experimental class during every period of teaching of the experimental unit. Most of the times the teacher took the major teaching role and the experimenter was his/her aide. A few times in each class, when the teacher felt particularly unprepared or insecure, the experimenter took over and switched roles with the teacher.

Classes were occasionally visited by outside observers who were invited by the experimenter to give some objective feedback.

2.4.4 Main lessons learned from the pilot study. Throughout the teaching period there was a natural tendency by the teachers to deal more with AC and DA cases than with MP and MT cases because the former caused more trouble whereas the latter were answered successfully most of the time.

Posttest item profile compared with pretest one (see appendix 7.3 for profiles) show a shift to "not enough clues" in wrong and right answers pattern. Results showed a small regression in the ability to apply MP and MT (see table 3.2, page 103) where most of the wrong answers in the posttest for these logical types were "not enough clues," instead of either yes or no on the pretest. This indicated an overlearning of the legitimacy of "not enough clues" as an answer.

The main lessons learned from the pilot study was that the first draft of the teachers manual was completely unclear and that there was a need for a more systematic and carefully planned pretraining workshop for the teachers.

As a result, the teachers manual was rewritten, integrated into the manual were many details about gimics and games pilot study teachers had used to attract their classes, dialogues to illustrate how to deal with wrong answers, and samples of charts successfully used by the teachers during the pilot study (see appendix 7.1). Objectives and administration procedures for each set of activities, including time planning suggestions, were also included in this manual.

CHAPTER 3

THE INSTRUMENT OF MEASUREMENT

Chapter Overview

Section 1.7.1 summarizes the objectives of this study, and sections 1.7.2 and 1.7.3 present the basic hypotheses. To measure the extent to which the objectives were fulfilled and to test the basic hypotheses, a test of 32 items was developed.

This chapter discusses the principal problems of developing the test (section 3.1) including difficulties faced in developing equivalent versions (section 3.1.3). It describes the final form of the instrument, its validity and reliability (sections 3.2.1 and 3.2.2). The last part of the chapter gives the procedures used to administer the test along with some precautions taken (sections 3.2.3 and 3.2.4).

3.1 Development of a Test

3.1.1 Limitations and relation with previously developed tests.

(a). In relevant literature one finds a distinction in the content of conditional reasoning problems between "causal relations" and "class inclusion" (O'Brien 1971, 1972; Roberge 1969, 1970; and section 1.4.6, page 39). Such a distinction is vague because any content in which causal relations are embedded can be expressed in terms of class inclusion. For example, even though the statement "if you enter the sea, you'll get wet" sounds completely causal, it can be interpreted in the following manner: "the set of events of your entering the sea is a subset of the set of events of your getting wet." In fact one often says, "Whenever you enter the sea, you get wet," which already has a class inclusion flavor. Since the distinction is not well defined, it was difficult to consider the distinction in constructing test items. Still in the test, about half the items have an embedded causal content, and half have a predominantly class-inclusion content (according to the subjective judgment of the researcher).

(b). All items are written in English and express content assumed to be familiar to fourth and fifth graders in California. This notion of "familiarity" gives rise to another ambiguity discussed in the literature, namely the distinction between factual content and a non-contradictory hypothetical content. Actually the two concepts differ only in the relative size of a relevant model that a reasonable student may readily have in mind for interpretation of the sentence. (The word "model" is used here in the sense used in logic, i.e., a universe in which the sentence becomes true.) To illustrate this point consider the sentence: "If the car is shiny, then it is fast", taken from O'Brien's studies. The collection of shiny, fast cars is one of many models for this sentence, because for this collection the sentence is certainly true. So if our children were living in a world where only shiny, fast cars existed, that sentence would be factual for them. However, for a Californian child this is not only a hypothetical sentence but also contradictory, in some sense, to his everyday experience with fast cars which are not at all shiny. It was not the purpose of this study to teach children to deal with hypothetical situations for which it is difficult to imagine a model or to test the extent to which they are able to do so. Therefore, items based on sentences like the one in the example above were omitted from the test.

The following is an example of a sentence which may be factual for some people but hypothetical for others. Consider the sentence: "If he has a driver's license, then his age is at least sixteen." It sounds factual to a Californian, but on second thought one may realize that this is not necessarily a fact for residents of some other country; other parts of the world may have lower age limits for obtaining a drivers license. Nevertheless, sentences such as the one above were considered to have familiar content and therefore as being legitimate in the test.

So the distinction between factual and hypothetical content is also not entirely clearcut. In the test, an effort was made to limit items to those made up of conditional sentences the content of which seemed (to the experimenter) to be reasonable for a nine to eleven year old Californian child. An effort was made to select items of content that were easy to picture, familiar, or described in previously learned terms, without distinction between a factual and hypothetical content. By no means would any item have a content contradictory to experience. The words "easy to picture" refer to sentences which a child of nine to eleven may have used before, or have heard in a natural conversation, and for which the child can easily devise a concrete model in which that conditional sentence holds. The next example is intended to illuminate this last point.

For children whose mothers are working it was difficult to accept the following sentence: "If George is sick today, then his mother will stay at home with him." Their first reaction was, "What if she has to go to work?" Still this sentence was not considered contradictory to experience even for those students, for it was considered that they could, without too much difficulty, imagine a suitable model. The same is true for: "If someone plays too much football, then he does not do enough homework." This sentence may raise emotional objection in some children; however, it was not regarded as one contradictory to experience, nor as one for which it was hard to imagine a model.

(c). It may happen that immediate experience removes the student's mind from a given item. For example, the puzzle: "If it is raining, then it is cloudy. It is not cloudy. Is it raining?", has a negative answer, which may contradict reality on a rainy day. All efforts were made to avoid items that could lead to such situations, but because we dealt with familiar content, this effort may not have been entirely successful.

(d). To build a list of items, conditional sentences were taken from previous studies (Hill 1961, Miller 1968, Carroll 1970) and modified or changed according to the limitations stated above. They were also carefully checked so that the content of each conditional sentence would not suggest its inverse (or converse). For example, Hill's (1961) sentence, "If Ann is at school, then she is the leader today," was omitted because the truth of its inverse -- "If Ann is not at school, then she is not the leader today" -- may be inferred from the original sentence, as

Ann could clearly not be the leader if she did not come to school. Since the original sentence functions as a biconditional sentence, it leads to a justifiable yes or no answer of AC or DA items built from this sentence, which would be counted as incorrect if the sentence is counted only as a conditional one. The same is true for the following example of Miller's (1968): "If Harry finds his meal ticket, then Harry can eat his lunch." This sentence, even though it does not logically imply its inverse, is very likely to suggest its inverse to many students who have experienced a lost lunch card, without having had money to purchase one, and who therefore had no lunch.

It is particularly difficult to design sentences that do not suggest their inverse in the "-+" negation mode (If not..., then ...), because in ordinary language this mode is usually used for dichotomious situations. For example, "If you do not feel all right, you should see your doctor." Obviously, if you are all right you do not need to see your doctor. Suppose, now, we include an item like this in the test

Clues: a. If you don't feel all right, then you should go see your doctor.

b. You feel all right.

Question: Should you go see your doctor?

A child using his common sense (along with logic) will answer -- no. For the child will see no need to see his doctor if he is all right (even though, in fact, he may need to see him for some other purpose). So the logical answer -- not enough clues - is unlikely to be used even by the good logical thinkers, because in addition to the given information they use some implicit knowledge that the content suggests.

The last example is also in the "-+" negation mode: If he is not here, then he is there." Now, if we are given that he is here, the common-sense answer to the question: "Is he there?" is "no." But, this conclusion does not follow logically from the first conditional sentence and the second given clue, but from the hidden information that no one can be in two different places at the same time. In other words the first conditional sentence is known to be a biconditional, because its inverse -- if he is here, then he is not there -- is always true. We can't expect children to ignore this knowledge and rely just on the logical validity of a conclusion, when their experience adds more information than is explicitly given in an item. It is in fact a general goal in education that, in the process of problem solving, a student will associate relevant knowledge and relate it to explicitly given data. It is surely not the purpose of this study to destroy this intuition. For this reason items such as the above were regarded as misleading and were not included in the test.

It should be noted, however, that as long as we do not deal with content-free items, problems of content are confounded with those of logic, and we can never totally avoid confusion caused by content, as in the above examples.

(e). As mentioned in section 1.3 (page 18), the immediate goal of the study did not involve quantificational logic, or sentences containing free variables. However, in ordinary language, very often the use of free variables implicitly quantified, provides the natural way of expressing a general rule. No effort to avoid such items was made. For example the following item was classified as MP, even though its second clue is not exactly the antecedent of its first clue but a particular case of it:

Clues: a. If a student does not finish the assignment, then the student has to stay after school.

b. Laura did not finish her assignment.

Question: Does she have to stay after school?

Finally, it should be admitted that criteria for including a certain item in the test were based more on intuition and careful examination of content than on definitions relating to the above ambiguities. Revisions of the initial test version were designed to increase the internal consistency of the test using an item-analysis technique which is described later on in this chapter.

3.1.2 Early versions and revisions. The test consists of 32 items. Each item contains two clues, one a conditional sentence, the other a statement, followed by a question that is never formulated with negation . Subjects mark their answer with an "X" to the left of the alternative they select from the answers provided: Yes, No, and Not enough clues. (See also notes about administration, section 3.2.3). The following is an example of an item:

Clues: a. If their car is not in their garage, then they are not home.

b. Their car is not in their garage.

Question: Are they home?

( ) Yes ( ) No ( ) Not enough clues

Notice that the questions are always stated positively, to avoid confusion which a negative question may raise. For example, if instead of asking "Are they home?" the question was "Are they not home?" some people would still answer "no," meaning "no, they are not home," namely using the no answer to reinforce the negation of the question rather than to negate it. In other words "yes, they are not at home" and "no, they are not at home" have the same meaning, even though from the strictly logical point of view a "no" answer to the negative question would mean "No. They are home."

In the format of an item, where the question is stated positively (see example above), the logic required to arrive at the right answer (no) is slightly more complicated than a simple application of modus ponens. For, after inferring by MP that they are not at home, one has to refer to the question -- Are they home? -- and give the answer: No. A second format was considered to purify the logic. In this second format the very same item would look like this:

Clues: a. If their car is not in their garage, then they are not home.
b. Their car is not in their garage.

Question: Which of the following sentences is therefore correct?

(1) Their car is in their garage.
(2) They are home.
(3) They are not home.
(4) Not enough clues to decide.

Despite the purer logic involved in answering the question in this format, it was found unsatisfactory for the following reasons: It requires much more reading; One of the alternative answers always contradicts the second clue (in this case the first one does); The answer is remote from the clues, making it harder to relate the clues to the conclusion; The phrasing on the whole is much more awkward than in the previous format. Later on, Paulus' (1967) format was considered. In his test he presents the items in the following way:

Suppose you know:

a. If their car is not in their garage, then they are not home.

b. Their car is not in their garage.

Then would this be true?

They are not home.

The objections to this format included ones similar to both objections to phrasing of the negative questions and objections to the second format. Therefore the first format (page 90) was used.

The 32 conditional sentences were chosen so that there were eight conditional sentences of each negation mode (++, +-, -+, --). In each negation mode, two items were then constructed for each logical form (MP, MT, AC, DA) by adding a second sentence that was either the antecedent or the consequent of the conditional sentence, or the negation of the antecedent or consequent. Thus, altogether eight items are obtained for each logical form and of these eight, two had the same negation mode. Such pairs, agreeing both in logical form and in negation mode, will be referred to as type-mates. (See table 3.4, page 108, for type-mate items.)

Each item is a three-choice question. Since for two of the four logical forms the correct answer is "not enough clues," the number of items correctly answered by "yes," "no," and "not enough clues" is eight, eight, and sixteen, respectively. The items were randomly ordered by blindly taking a numbered card from an urn of 32 numbered cards, without replacement. If four consecutive items chosen this way had the same logical form, the fourth one was put back and a new card drawn.

A revision of the first version of the test was made after administering it to 42 fourth graders in Albany, California, during April 1974. The revision was based upon analysis of item-scores within logical forms. Any item that scored inconsistently with its type-mate, or whose scoring seemed exceptionally different from all other items of the same logical form, was changed or replaced to increase internal consistency within logical forms, still keeping in mind the expected effect of negation mode. Reading difficulties, such as in reading the word "judo" or the name "Marian," were discovered through students' request for help, and also led to modification of test items. To separate the question itself from the clues given as the basis for the answer, the two clues were numbered and the word "Question" was added before the question. The need for such change was realized through administration of the test.

Results of the test showed that students very rarely used "Not enough clues" as an answer to any of the questions. It was felt that they might be reluctant to use it because they interpreted this answer as an admission of their own personal inability to give an answer rather than as an assertion of the logical indeterminability of a definite answer to the question. It was hypothesized that this confusion might be reduced by adding examples in the introduction of the test.

To test this hypothesis two revised versions of the test were administered to a second group of 30 fourth-graders in Albany, California, in late April 1974. The two versions were identical except for their introduction. In one version there was a description of item structure and answering procedure with no example. In the other version, in addition to that description, two examples were given, one of which has "No" as its right answer, while for the other one "Not enough clues" was right (see appendix 7.2). A brief explanation was given for each answer to the introductory questions. Students were randomly assigned to the two versions. The analysis of the results showed that students who took the version with examples were less reluctant to use "Not enough clues" as an answer. On the average they used it in $\frac{1}{5}$ of the cases (either rightly or wrongly) compared to only in $\frac{1}{16}$ of the cases for the group who took the version with no introductory examples. Total test scores were higher (however, not significantly so) for the version with examples than for the version without examples. Introductory examples, therefore, seemed to give a better explanation of the answer "Not enough clues" than just a verbal description of the possibility of that answer applying. However, the need for further learning of the logic behind the use of that answer remains. This improvement of the introduction may increase pretest scores, and thereby avoid inflated pretest - posttest gain scores unduly attributed to the treatment effect.

It should be kept in mind that questions without a definite answer occur very rarely in the regular school curricula, and so learning that "not enough clues" is a legitimate answer is an achievement which should not be underestimated. Obviously, however, the purpose of this study goes much further. Beyond the learning that there may be insufficient information to answer a given question, the study was designed to improve the ability to use logical analysis.*

The form of the test with examples in the introduction was used in all experimenter-conducted teaching phases of the development of the experimental unit, described in Chapter 2. In all the trials, administration of the test lasted no more than 35 minutes, including at most ten minutes for the introduction. This was considered reasonable for the attention span of students at that age, and consequently no change in the length of the test was made.

The next step in the development of the test was to obtain equivalent versions for use as pretest and posttest. This effort presented a formidable problem.

3.1.3 From one test to four equivalent tests. The study was originally planned as a time-series study. Each student was to be tested four times and introduced to the experimental unit between the second and third tests. This design was later simplified, mainly because experience with just pre- and post-testing led to the suspicion that taking similar tests repeatedly four times would significantly decrease student interest. In consequence, the number of random answers in the third and fourth tests, along with an effect of learning just from the tests, would jeopardize the results of the experimental teaching.

---

*For a discussion of the extent to which this was achieved, see section 5.5, page 160.

Another reason for limiting the final design to two tests was the difficulty experienced by the experimenter in developing four equivalent tests. The following is a description of this effort:

(a). Equivalence based on negation modes. Given any two statements, p and q, one can construct four conditional statements, different from each other in their negation modes, by negating either p or q or both. In this way one gets "if p then q;" "if p then not-q;" if not-p then q;" "if not-p then not-q." It would appear natural to use this fact to obtain four equivalent tests by constructing an item of identical logical form for each of the four negation modes and putting one of them in each of the four tests. For example, by adding a second clue which matches the antecedent of each of the four negation modes, one gets four different MP items. To obtain the right combination of negation modes within logical types in each test, one could then permutate the order of assigning items to tests. Even though this approach appears to be very clear and elegant, it involves tremendous complications when meaningful content is considered. It is extremely difficult to find conditional sentences that are meaningful in all four negation modes. It happens often that only two of the four negation modes make sense, usually the two are either $p \rightarrow q$, $-p \rightarrow -q$ or $p \rightarrow -q$, $-p \rightarrow q$.*

---

*See also discussion of contrapositive transformation later on in this section for further difficulties.

Even if 32 conditional sentences were formed which did make sense, and did obey all the limitations discussed in section 3.1.1 and 3.1.2, obviously the four tests obtained by this method would not have item meaning preserved. Because the intent was to come as close as possible to a content-free measure, so as to separate out content influence, a set of four tests having this deficiency would be unacceptable, and deemed not to be equivalent forms, even before giving them a field trial. Even though as measuring instruments each of them may be valid, because each is a direct measure of the objectives, their results would not be considered comparable.

(b). Equivalence based on logical forms. Because operating on the conditional sentences was found unsatisfactory as a way of producing variant questions, operating on the second clue of each item was considered. Here again there is a natural way to obtain four different items by constructing four items of different logical types based on a single conditional sentence as a first clue. The method is to add as a second clue either its antecedent, or its consequent, or the negation of one of these. One of each set of four items obtained in this way would then be assigned to each of the four different tests. Each of the four tests would then consist of 32 items which are identical in their first clues (i.e., the conditional sentence). However, this should not be misinterpreted as meaning that the four tests are item-equivalent. The logical considerations which enter an MP or MT logical form item are completely different from those entering AC and DA items, even where they are all based on one conditional sentence. It is true that an application of MT is nothing but MP applied to the contrapositive, just as MP is an application of MT to the contrapositive. It is also true that AC and DA are similarly related. However, there is no way by which to relate MP (or MT) to AC (or DA). True, both MP and AC are affirmative in character and both MT and DA have denial properties, but these features are logically irrelevant, and must be considered artificial relations which may, at most, have something to do with Sells' atmosphere effect.* It is unreasonable to assume equivalence of items paired on this basis. So again, even before considering a field trial the use of four such tests was rejected for lack of grounds to hypothesize their equivalence.

(c). Mixed order by type-mate interchange. As was previously mentioned, the test consists of 32 randomly ordered items in 16 pairs of itentical type, where type-mates have the same logical form and negation mode. The easiest and obviously the safest way to obtain an equivalent version of the test would be by interchanging the place of type-mates. This procedure preserves the order of right answers as well as the order of item types, and of course it preserves item-by-item content. If the original form had a high internal consistency, the mixed order version has a high chance of preserving split-half reliability provided the two halves consist of one item of each pair of type-mates. This is because the two halves would be identical for the original and the mixed order version. Also, any deviation from a perfect test/re-test reliability, using the other version in the retest, must in this case be attributed to an error external to the test itself.

---

*See section 1.4.4.

The possibility of students memorizing pretest items was considered. This problem was determined to be insignificant because of the large number of items and the big difference that a change in one word would make.

For a given test T, T' will denote a mixed order version. The mixed order transformation reduced the need for the number of equivalent tests from four to two. Out of each test, the mixed-order version would then be obtained, and altogether there would again be four.

(d). Transformation by contrapositives. Given a test of 32 items of the structure described in section 3.1.2 (page 90) one can express the conditional sentence (the first clue) of each item in the contrapositive way namely, "if a then b" will become "if not b, then not a" , and keep the second clue and the question unchanged. For example, an AC item may originally say

Clues: a. If it is rainy, then it is cloudy.
b. It is not rainy.
Question: Is it cloudy?

Its image under contrapositive transformation will then be the DA item saying:

Clues: a. If it is not cloudy, then it is not rainy.
b. It is not rainy.
Question: Is it cloudy?

This transformation obviously changes the amount and place of negation in any item, but formally preserves the content, at least as far as its truth value is concerned. Therefore the right answers are preserved also. Practically, however, formulation of contrapositives leads to a severe grammatical problem with tenses, for English grammar requires no future tense in the antecedent. Most often this makes the contrapositive expression very awkward and less clear. E.g., "If you don't clean your room, you'll not get your allowance," will be transformed into "If you get your allowance, you'll clean your room." It may also be noticed that this transformation does not preserve the logical-form of items because it interchanges the logical forms MP and MT, as well as the forms AC and DA. To avoid these grammatical difficulties, the idea of transforming just logical types was reconsidered. This idea is discussed below.

(e). Transformation of logical types. Part (b) of this section discussed the difficulties involved in designing four equivalent tests based on systematic variation of the logical form. Since only two versions were now needed, two possible transformations on the logical type were reconsidered. In both versions the conditional sentence of the transformed item remains unchanged, and the logical type is changed by operating on the second clue, and by modifying the question accordingly. These transformations accomplish the following interchanges:

(i) MP ↔ AC; MT ↔ DA

(ii) MP ↔ MT; AC ↔ DA

In (i) an item with a definite answer is changed into one with an indefinite answer item and vice versa but the affirmative characteristic (or the denial characteristic) of the second clue is kept, and in transformation (ii) it is the other way around. Arguments similar to those which appeared in part (b) of this discussion concerning methods for creating four equivalent tests based on constructing four items of different logical types from one conditional sentence, led to a preference for transformation (ii). Using this, a second version of the test was obtained. These versions will be referred to as T_1, T_2. A copy of each can be found in Appendix 7.3. T_1, T_2 were tried out in the pilot study with the hope that they would prove equivalent; then, by type-mate order interchange of each, four forms would be obtained. The extent to which this hope was realized will be discussed in the next section.

3.1.4 Field trial of two equivalent versions. The two versions T_1, T_2 mentioned in the last section (see also appendix 7.3) were used for measurement of students' progress in the pilot study. The pilot study itself was described in section 2.2.4. In each class students were randomly assigned to the test versions by handing out the same version to every other student so that two neighbors always received different versions. Those students who had T_1 for their pretest got T_2 in the posttest, and vice versa. By this method each class was randomly halved. The two halves will be referred to as T_1-T_2, and T_2-T_1, according to the pretest - posttest versions order. Had the two halves performed equally well on the pretest and equally well on the posttest in both treatment groups (those who got instruction using the experimental unit between tests), and in the control group (where no exposure to the experimental unit was given between tests), the equivalence of the two forms and the reliability of each as a measurement would have been established, because this would have shown equal pretest to posttest gain score no matter which version was used. Such a validation procedure is based upon the assumption that the two halves in each class are equivalent due to the randomization. It turned out, however, that this assumption was not quite realized -- at least not as far as math and reading achievement scores were concerned. The following table shows the students' level of achievement in these regular school subjects as determined by the results on the CTBS tests taken prior to the study.*

Table 3.1 Percentage of Pilot Study Subjects in CTBS Reading and Math Levels

Group	Level	Reading		Math	
		T_1-T_2	T_2-T_1	T_1-T_2	T_2-T_1
Experimental $n(T_1-T_2) = 24$ $n(T_2-T_1) = 25$	High	37.5	48.0	37.5	31.6
	Medium	50.0	28.0	45.8	47.4
	Low	12.5	24.0	16.7	21.0
Control $n(T_1-T_2) = 38$ $n(T_2-T_1) = 37$	High	34.2	27.1	31.6	37.8
	Medium	34.2	32.4	47.4	29.8
	Low	31.6	40.5	21.0	32.4

The above data would not refute the assumption of equivalence of the two tests, if performance on the tests showed a great similarity between the T_1-T_2 and T_2-T_1 scores. They would, however, rule out the possibility of validly drawing any conclusion about the equivalence of the two versions, in case the T_1-T_2 and T_2-T_1 scores came out rather different from each other. Table 3.2 gives

---

*CTBS is the Comprehensive Test of Basic Skills, see bibliography.

Table 3.2 Pilot Study Mean Scores by Versions Order

Group and version order / Logical form		EXPERIMENTAL		CONTROL	
		T_1-T_2 n= 25	T_2-T_1 n= 24	T_1-T_2 n = 38	T_2-T_1 n = 37
MP	Pre	6.4	6.1	6.3	5.8
	Post	5.7	5.2	6.0	5.8
	Gain	- .7	- .9	- .3	.0
MT	Pre	5.6	5.5	5.7	5.7
	Post	4.9	5.1	5.8	5.8
	Gain	- .7	- .4	.1	.1
AC	Pre	1.9	1.7	1.4	1.7
	Post	4.5	5.2	2.2	2.2
	Gain	2.6	3.5	.8	.5
DA	Pre	2.5	1.7	1.7	1.6
	Post	5.1	4.9	1.7	2.8
	Gain	2.6	3.2	.0	1.2
Total	Pre	14.8	15.0	15.6	14.7
	Post	20.3	20.5	15.7	16.6
	Gain	5.5	5.5	.1	1.9

pretest, posttest, and mean gain scores of the total test scores, as well as the subtests scores, according to T_1-T_2/T_2-T_1 for the experimental and control groups of fourth grade classes participating in the pilot study.

As follows from the discussion above, and from table 3.2, equivalence of T_1 and T_2 was neither established nor rejected by the pilot study. This is because on some parts performances of the two group-halves were not equal for either the experimental or the control group. Consequently it was decided to use a revised T_1 and its mixed order version (see section 3.1.3) as the instrument for the principal study.

3.1.5 Final revision. Item profiles for each of the 32 items of T_1 were prepared by counting the number of yes/no/not-enough-clues answers given on the pretest by T_1-T_2subjects, and on the posttest by T_2-T_1 subjects. The profiles themselves are given in appendix 7.3. Within each logical form, profiles of type-mate itmes were compared with one another both for the experimental and the control groups. When an inconsistency was found, the pair of items was reviewed, and modified wherever this review yielded a reasonable explanation for the gap. These modifications will be discussed in detail in the rest of this section up to the last paragraph, where modifications of the test outlook are described .

Table 3.3 Final Revision of the Test

Item	was	changed to	Reason for making the change
3	If the sun is not shining, then...	If the weather is not warm, then...	Makes the content more realistic
3	Will Cindy go swimming?	Is Cindy going to swim?	Content interfered with logic. We can't tell whether in the future Cindy will or will not go swimming, because weather may change.
5	If Laura's desk is not straightened up...	If Sue's desk is not cleaned up...	Reading difficulties discovered through subjects' questions.
8	If there is a holiday next Wednesday, then the library will not be open. The library will be open next Wednesday. Is there a holiday next Wednesday? (No)	If it is a holiday, then the library is not open. The library is open. Is it a holiday? (No)	Content interfered with reality. There may be a holiday on the Wednesday following the test.

(continue)

Item	was	changed to	Reason for making the change
9	If the wind does not change direction, then...	If the wind does not change, then...	Better clarity by shortening and simplifying the sentence.
11	If Jack is not in the race, then Joe's team will win. Jack is not in the race. Will Joe's team win? (Yes)	If Jack is not in the race, then his team will win. Jack is not in the race. Will his team win? (Yes)	Need only think about one team instead of two. Also, content fits reality better this way.
13	...Janet does not come home in time. Are her parents worried?	...Janet did not come home in time. Were her parents worried?	Better English in making second clue a particular case rather than a general one.
14,27	Interchange of the first clue between these two items was made in order for the content to better fit the logical form of the answer.		
14	If that woman is Mrs. Brown, then she is Nancy's grandma. That woman is Nancy's grandma. Is she Mrs. Brown? (NEC)	If the aquarium is dirty, then the goldfish will die. The goldfish has died. Was the aquarium dirty? (NEC)	Mrs. Brown is one of just two grandmothers. This small universe may cause trouble. Since one grandma is mentioned, the existence of another may be overlooked. In the aquarium case, children are aware of many reasons for death other than dirt.
18	If a record has no crack in it, then it is not Jeremy's.	If a record has no crack, then it is not John's.	Difficulties in reading the name Jeremy plus shortening by omitting the redundant words "in it."
22	This item was one of the only two items replaced (see 26 below). Its content is contradictory to experience; every puzzle fan knows that not all little pieces fit any puzzle. Given in MT form, the item is even more unacceptable. Its contrapositive originally appears in Hill's study.		
26	This is the second item replaced (see 22 above). The word mammal caused a lot of trouble to many subjects. This unfamiliar word made the item seem more of a nonsense syllable item than the factual-content item it was intended to be.		

(continue)

Item	was	changed to	Reason for making the change
27	If the aquarium is not clean, then the goldfish will die. The goldfish has died. Was the aquarium clean? (NEC) (Also see note on this item preceding item 14.)	If that woman is not Mrs. Brown, then she is Nancy's grandma. This woman is Nancy's grandma. Is she Mrs. Brown? (NEC)	It is very tempting to answer NO to the original item. Even if there are many other reasons for the goldfish's death, the aquarium was probably not clean too. Compare the original item here with the modified item no. 14. The change from not clean to dirty seems to make a lot of difference in one's readiness to think of other reasons for the goldfish's death.
29	If I don't see him today, then I'll see him tomorrow.	If I don't see Dennis today, then I'll see him tomorrow.	The word "him" occurs too many times in the original item. An unnatural way of using a free variable.
30	If this house has a red roof, then it is not Joy's house.	If a house has a red roof, then it is not Joy's.	Better English and shorter sentence.
32	If he takes music class, then he is not supposed to be here.	If a students takes Spanish, then he does not take French.	Pattern of wrong answers was inconsistent with previous studies (O'Brien) and with other DA items. No reason was hypothesized except content making no sense to students. Hence content was changed.

In general, the modifications were made to clarify items and to avoid interference of content with logic. The modifications were supposed to increase pretest scores in the main study, which in fact were higher, though probably not only due to test modifications. In this way the validity of interpreting gain scores as a true learning effect would be increased.

Finally, some "outside" changes in format were made. Each item was typed on a separate page, the page size was reduced to half of a regular-size page, and the word "not" was underlined wherever it belonged to the logical form, i.e., in MT and in DA forms. (It was not underlined when it belonged to the negation mode, i.e., in MP and AC forms of +-, -+, or -- negation modes.) Minor changes in the wording of the instructions may also be noticed by comparing the pilot-study version of T_1 (appendix 7.3) with the final version (appendix 7.2).

3.2 The Final Form of the Instrument

3.2.1 The two versions. The instrument consists of two versions of a 32 item test. In each version there are 16 pairs of items of the same type (negation mode and logical form). Table 3.4 shows item numbers classified by type. The right answer in each case is shown in parentheses. The two versions (denoted by T and T') differ only in the order of type-mates. Namely, item 1 in T is identical to item 23 in T', and item 23 in T is identical to item 1 in T'; similarly for each pair of type-mates. The test itself is given in appendix 7.2.

3.2.2 Validity and Reliability. Except for its particular content, the test is a direct measure of the ability to validly infer from conditional sentences. As such, its validity is by definition unquestionable. The fact that content is sometimes confounded with logic does not reduce its validity at all. In fact it is the other way around. Since it is difficult to imagine any content-

Table 3.4 Test Item-Numbers by Type

Negation mode	Logical Form			
	MP	MT	AC	DA
++	1,23 (yes)	22,31 (no)	2,14 (NEC)	17,10 (NEC)
+-	25,30 (no)	8,16 (no)	4,7 (NEC)	6,32 (NEC)
-+	11,13 (yes)	5,9 (yes)	21,27 (NEC)	15,29 (NEC)
--	3,19 (no)	20,28 (yes)	12,18 (NEC)	24,26 (NEC)

free application of logic in a "real life" setting, the variety of content which appears in the test strengthens its content-wise generalizability within the limitations discussed earlier (see section 3.1.1).

Since reliability is regarded as a necessary condition for validity (Anastasi 1968), the reliability of the test follows. However, reliability has many faces. It is not uniquely defined, and hence its necessity for validity is not in the strict "by definition" mathematical sense. Because of such doubts, an independent reliability study was carried out in two ways, as follows.

(i) A test/re-test Pearson product moment correlation was computed on pretest/posttest total scores in the control group of the principal study. This gave a reliability of .75.

(ii) A split half Pearson product moment correlation coefficient with Spearman-Brown formula for double length was computed on pretest total scores for both experimental and control group subjects. This gave a reliability of .79.

For an analysis of item profiles and discussion of internal consistency (a third point of view on reliability) see section 5.1.2, page 126.

3.2.3 Administration. The five front pages of the test (see appendix 7.2) were duplicated on colored paper: yellow for T and green for T'. The experimenter was introduced to the class by the teacher as a friend who came from the Lawrence Hall of Science, a science education center at the University of California, which the students had visited on field trips. The teacher told the class that the friend had come to give them some puzzles for fun. Page two of the test was then read by the experimenter; in fact, it was almost known by heart, so that eye contact with the class was continuous. When finished reading page 2, the experimenter handed out the puzzle books. Students who sat next to each other were assigned to opposite colored teams (different versions). Students were asked to fill in the blanks on the front page and, to make the contest fair, not to turn this page until they were asked to do so. After everyone had finished the front page, students were asked to turn to page 3 and to follow the experimenter's reading. When examples on pages 4 and 5 (see appendix 7.2) were given, students first answered, and then the right answer was read along with its explanation. The experimenter circulated to make sure students knew how to mark their answers.

This introduction took 8-10 minutes in each class. Within the next 20-25 minutes students were working on their own and the experimenter circulated to answer the very few (one or two in each class) requests for help in reading a word.

The above mode of administration was followed on the posttest, except for changing each team's "color" (version), and changing seats where necessary. Of course the introduction of the experimenter by the teacher in the control group was changed -- "The lady from the Lawrence Hall of Science is here again for another period of puzzles for fun" -- and the experimenter introduced herself to the experimental group, saying: "You've learned a lot in logic. Let's see how well you can do this time in the puzzle books."

3.2.4 Precautions taken in administration of the test. The test was presented to the students as a green versus yellow team contest in puzzle solving. The teams were the two halves of the class created by handing out different versions to neighboring students. This was intended to serve two purposes:

(i) Increase student motivation to take the test, and in this way reduce random or careless answers by creating a competitive collective atmosphere, and

(ii) Insure independent work by each student. Due to the manner of distributing the two versions, every student was surrounded by students of the opposite team.

In order to avoid observation of incidental effects, another precaution was taken to increase student chances for success in the test, even before any training. The instructions given prior to taking the test include two examples, one of which has "Not enough clues" as the right answer. In addition a verbal explanation of the "not-enough-clues" alternative was given.*

---

*See page 5 of the test, appendix 7.2, and previous discussion of this point in section 3.1.2, page 94.

It should also be noted that a no-time-limit announcement was made, and help in reading was offered when necessary.*

To decrease the amount of pure guessing, students were told they would lose points for wrong answers (see page 2 of the test, appendix 7.2). They were not penalized for an unanswered item. However, since there was no time pressure, very few students left questions unanswered, and those who did skipped at most two, but mostly one, item. (See exact numbers in table 5.3.) A special note in the instruction prior to the test was designed to help the students get organized, by suggesting the use of a special mark next to a skipped item. This also served as a way to check that a student who skipped an item did it intentionally, and not just by missing a page.

To prevent peer pressure of faster ones on slower ones towards the end of the testing period, tests were not collected until all the students in the class had finished their work. Those who finished early were, if necessary, referred back to page 2 of the instruction where they were asked to draw a picture or a design on the back of any page. (Many were reluctant to stop drawing when the test was over....)

Team scores were given to the classes a few days after each test. Individuals' scores were given to teachers. No correction period followed and puzzle books were not returned to students. Students seemed to enjoy the pretest a great deal, but were less enthusiastic, even though cooperative, the second time (posttest). This change of attitude was particularly noticeable in the experimental group.

---

*See page 2 of the test, appendix 7.2.

CHAPTER 4

THE PRINCIPAL STUDY

Chapter Overview

This chapter outlines the main study. Section 4.1 lays out the experimental design and discusses the reasons this particular design was selected. Specific data on the population sample are provided in section 4.2. Section 4.3 details the course of the main study. This last section also provides information on the teachers, their training, and their subsequent teaching.

4.1 Research Design

4.1.1 Lay-out of the design. The main study applied Campbell and Stanley's (1963) quasi-experimental design no. 10. Basically this design is a pretest-posttest, treatment - no treatment control group design. Experimental subjects are not assigned randomly, from a common population, to the experimental group and the control group, so the two groups lack a pre-experimental sampling equivalence.*

In the present study, the experimental and the control groups each consisted of the entire fifth grade, each in one school, composed of four classrooms with their regular teachers.

Both groups were drawn from the same school district, which has a relatively homogeneous population. Diagram 4.1 illustrates the design.

---

*Refer to section 4.1.2 for reasons this design was selected.

Diagram 4.1 Research Design

T, T' refer to the versions of the test (see section 3.2.1 and appendix 7.2);
X_i (i=1,2,3,4) stand for the four characteristic styles of teaching the experimental unit by the individual teachers;
Y_i (i=1,2,3,4) stand for the teaching in the control classes, which was likely to be irrelevant to the experimental unit. The parentheses in the diagram indicate this irrelevance;
The dashed lines indicate a non-random assignment.

Experimental group (school 1)	class 1	T	X_1	T'
		T'	X_1	T
	class 2	T	X_2	T'
		T'	X_2	T
	class 3	T	X_3	T'
		T'	X_3	T
	class 4	T	X_4	T'
		T'	X_4	T
Control group (school 2)	class 1	T	(Y_1)	T'
		T'	(Y_1)	T
	class 2	T	(Y_2)	T'
		T'	(Y_2)	T
	class 3	T	(Y_3)	T'
		T'	(Y_3)	T
	class 4	T	(Y_4)	T'
		T'	(Y_4)	T

The experimental group and the control group were each pre-tested and post-tested by the experimenter in their regular classes, on a single day. The experimental group was pre-tested one day before the control group, and post-tested one day after.

Control group teachers were not pretrained or provided with the experimental materials.

4.1.2 Why a quasi-experimental design? The quasi-experimental design described in section 4.1.1 was selected because it allowed for the entire class to work as a whole in as normal a mode as possible. This seemed to have many advantages over the corresponding true experimental design in which each class would be randomly split into experimental and control groups. The advantages of this quasi-experimental design follow:

a. This design most closely resembled real classroom teaching which was one of the objectives of the study. There was, for example, no need to change rooms or any classroom organization for carrying on the study.

b. The design avoided to a large extent students' awareness of participating in an experiment (the "I'm-a-guinea-pig" attitude, Hawthorne effect, etc.).

c. Classes were taught by their regular teachers. Application of the unit by ordinary classroom teachers was another objective of the present study. This factor also may have contributed to a "business as usual" atmosphere.

d. The design avoids much of the communication among experimental- and control-group subjects concerning the experimental unit. In the true experimental design such communication is likely because classroom peers would be assigned to different groups for the research purposes. Then they would meet again for regular school work. Curiosity on the part of the control-group subjects, and pride or showing off on the part of the experimental-group students would increase the likelihood of such communication.

e. Preliminary investigations (see section 1.4) suggested that pretest scores would show similarity between the experimental and the control groups on the ability to distinguish between valid and fallacious inferences in conditional logic.

f. This design, regardless of its quasi-experimental nature, to use Campbell and Stanley's language:

> "should be recognized as well worth using in many instances..., The more similar the experimental and the control groups are in their recruitment, and the more this similarity is confirmed by the scores on the pretest, the more effective this control becomes. Assuming that these desiderata are approximated for purposes of internal validity, we can regard the design as controlling the main effects of history, maturation, testing and instrumentation, in that the difference for the experimental group between pretest and posttest (if greater than that for the control group) cannot be explained by main effects of these variables such as would be found affecting both the experimental and the control groups." (Campbell and Stanley, 1963, pages 47-48.)

4.2 The Sample

All four fifth grade classes in Los Perales School at Moraga School District, California, served as the experimental group. (They will be referred to as E_1, E_2, E_3, E_4, where E stands for experimental.) In the same school district, Rheem School's four fifth grade classes served as control classes. (They will be referred to as C_1, C_2, C_3, C_4; C stands for control.) Moraga is a suburban community in the San Francisco Bay Area. The socioeconomic categorization of the families within this school district is predominantly upper-middle class. There is no tracking in this grade. Table 4.1 gives basic data on the subjects.

Table 4.1 Basic Data on the Sample

	Experimental group	Control group
Total number*	104	106
Males	49 (47.1%)	48 (45.3%)
Females	55 (52.9%)	58 (54.7%)
Age (in months): mean	130.9	129.9
SD	4.8	4.6

The Stanford Achievement Test (1973 edition) was administered in all eight classes in October 1974. Table 4.2 gives the mean national grade-equivalent and national stanine scores summary for each class participating in the study.**

---

*Only those 210 students who participated in both pretest and post-test and who were not absent for more than five days during the learning period were included in the analysis.

**D. A. Payne (1968) uses the following figure to illustrate the relationships between the normal curve, the stanine standard scores, and IQ (Payne, 1968, page 111):

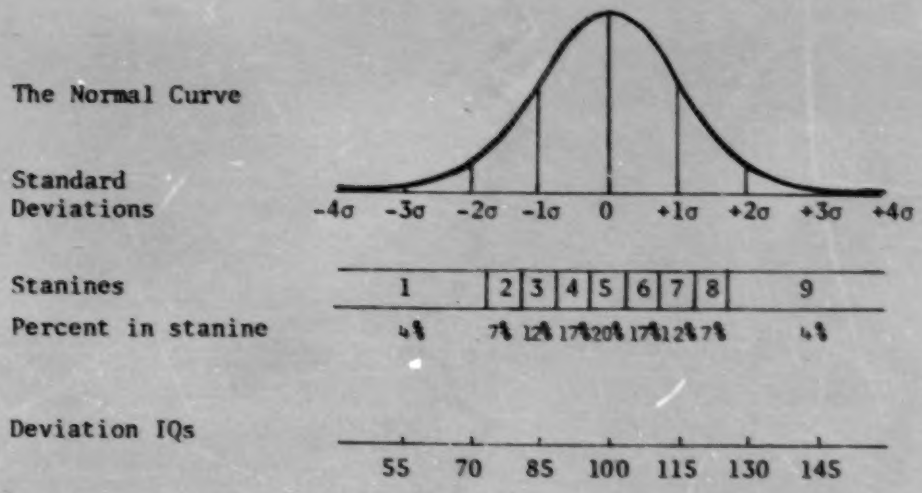

The Normal Curve
Standard Deviations
-4σ
-3σ
-2σ
-1σ
0
+1σ
+2σ
+3σ
+4σ
Stanines
1
2
3
4
5
6
7
8
9
Percent in stanine
4%
7%
12%
17%
20%
17%
12%
7%
4%
Deviation IQs
55
70
85
100
115
130
145

Table 4.2 Stanford Achievement Test Scores for Experimental (E) and Control (C) Classes

Class	N	Mean National Grade Equivalent	National Stanine Summary					
			Below Average (1,2,3)		Average (4,5,6)		Above Average (7,8,9)	
			N	%	N	%	N	%
E_1	26	6.7	0	.0	12	46.2	14	53.8
E_2	25	6.5	1	4.0	11	44.0	13	52.0
E_3	28	6.8	1	3.6	12	42.9	15	53.6
E_4	25	5.9	2	8.0	15	60.0	8	32.0
C_1	26	6.3	1	3.8	11	42.3	14	53.8
C_2	27	6.6	1	3.7	14	51.9	12	44.4
C_3	25	6.7	0	.0	10	40.0	15	60.0
C_4	28	5.9	1	3.6	17	60.7	10	35.7

4.3 Course of the Project

4.3.1 First Contacts. In late November 1974 three superintendents of school districts in the San Francisco Bay Area were contacted through their curriculum specialists. Two of the three agreed to meet with the experimenter for a discussion of the proposed project and a presentation of sample materials. After these meetings, each superintendent arranged for the experimenter to meet with fourth and fifth grade teachers. Ten to twelve teachers in each school district attended the meetings. In these meetings, which occurred in mid-December 1974, the experimenter provided the background for the project and described the project's goal. The experimenter also demonstrated a few activities and discussed the experience gained in the pilot study.

The last part of each meeting was devoted to the level of commitment expected from those teachers who agreed to have their classes participate in the study. This commitment consisted of a pretraining workshop of 9-12 hours in 6-9 sessions during January 1975, a once-a-week staff meeting during the teaching period, employment of the project during February 1975, evaluation of the activities during the implementation, and an overall evaluation paper at the end. Teachers were informed that the University extension of the University of California at Berkeley was expected to approve their project work for an in-service course with 3 quarter units of credit. Both school districts agreed to grant the teachers the equivalent credit if the University did not approve the course. (The course was approved by the University in February 1975 and four participating teachers received credit.) The individual teachers had until January 1, 1975, to decide whether they wanted to participate in the study. All fifth grade teachers in Los Perales School in Moraga School District volunteered and fully participated in the project. From here on the term experimental group will be used for the four classes in Moraga. (E_i, i = 1,2,3,4 will stand for these experimental classes.) The four control classes were obtained with the assistance of the district office.

4.3.2 The teachers. Table 4.3 gives basic information on the teachers and their background. The teachers of classes E_3 and E_4 team-teach throughout the year. In general all four teachers of the experimental group have good communication and regularly share their plans and experience.

Table 4.3 Basic Data on the Teachers

Class	Sex	Years of Teaching Experience	A.B. Degree & Credential			Previous Logic Courses	Grad. Work Beyond Credential
			Year	Major	University		
E_1	M	2	1972	Humanities	UC Berkeley	none	none
E_2	M	1	1969	Political Science	UC Berkeley	none	some
E_3	F	3	1969	History	Vassar & UC Berkeley	philosophy	none
E_4	F	1	1973	Elementary Education	Pacific U. Oregon	none	none

4.3.3 The pre-training workshop. During the three weeks from January 13th to 31st, six sessions of two hours each were held twice a week after school in Los Perales School. Teachers were instructed in the logic of conditional reasoning with the use of the experimental unit materials. The experimenter presented the activities to the teachers the way the teachers were expected to present the activities to their classes. Because the teachers had no background in logic either, they easily played the student role.

Teachers were warned not to overemphasize AC and DA cases, and were asked to keep a good balance of all four logical forms in order to prevent the regression in the MP, MT cases. (Eventually, regression in them was negligible in the main study.)

It was repeatedly stressed that the purpose of the experimental unit was not to teach the algorithm.* If some students discovered it, fine. If others did not - that would also be fine.

---

*See algorithmic solution in section 1.3.8, page 23.

In addition to the experimental materials presented, a broader view of the subject was also introduced as part of the general orientation for the teachers. In each session the teachers received the appropriate part of the teacher's manual. They were assigned reading and homework for the next session. A session-by-session account of the training workshop is given in appendix 7.5, page 392.

During the workshop teachers were appreciative of the importance of trying out the ideas incorporated into the experimental unit. They were willing to make the effort to provide detailed feedback to the experimenter. They brought up ways of raising students' motivation and thought about means for eliminating frustration.*

4.3.4 The teaching period. The four experimental classes, which constituted the experimental group for this study, were pretested by the experimenter on February 5, 1975, and posstested on March 21, 1975. Advance notice was given only before the posttest.** Between the two tests the experimental-group teachers taught the experimental unit four to five periods of 30-40 minutes each week. February 12-17 were vacation days so altogether each teacher spent 23-25 sessions with his/her class working on the experimental unit.

Once a week after school, the experimental-group teachers met with the experimenter to discuss the previous week's experience, the following week's plans (including practical considerations, e.g., grouping of students), current teaching problems (e.g., motivation), handling of students' difficulties, and review of the underlying logic.

---

*See section 5.8 for a discussion of teachers' attitudes throughout the implementation of the experimental unit.

**Control group was tested on February 4 and March 20, 1975. No advance notice was given for either test.

The experimenter visited each experimental class only twice in the period between the two tests in order to maintain a low profile. The main purposes of these visits were to see that teachers did not teach the algorithm, and to hear and record students' arguments. There was no evidence that any teacher released the algorithm, but more than a few students in each class, according to teachers' reports, sensed quite rapidly the existence of a pattern. In each class, two to three students were able to express this pattern in terms of the relation between the second clue and the parts of the first clue. Results of the study are reported and analyzed in the next chapter.

CHAPTER 5

RESULTS AND THEIR INTERPRETATION

Chapter Overview

In the following pages an analysis of pretest, posttest, and gain scores is given. The pretest equivalence of experimental and control groups is established in section 5.1. Progress of the two groups is compared in sections 5.2 and 5.3. The groups findings show a significant difference between the posttest scores of the experimental group and the control group on the undecidable logical forms AC and DA, and no significant difference on MP and MT logical forms on which both groups performed successfully in the pretest, already.

Analysis of the negation-modes subtests is in section 5.3.4. It leads to a discussion of the false conclusions that can mistakenly be drawn by overlooking the differences of negation-mode order of difficulty within logical forms. In section 5.4 an analysis of variance on gain scores is reported.

An analysis of right and wrong answers appears in section 5.5. This analysis attempts to sort out the guessing, particularly on the "not-enough-clues" answers.

Section 5.6 contains a correlation study of the Stanford Achievement Test results and results of the present work. The chapter closes with a discussion of student attitudes in section 5.7. The group was almost equally divided among those who enjoyed the unit and those who did not. The teachers' evaluation of the unit and their changing attitudes are discussed in detail. Included also are teachers' criticisms of the unit's length and of the repetitive nature of the activities in the unit. Also expressed is the teachers' appreciation of the need for such a unit and their willingness to use it again in a modified version next year.

5.1 Pretest Scores

5.1.1 Equivalence of control and experimental groups. Because the assignment of students to experimental and control groups was not random,* the first point analyzed was the extent to which the two groups were comparable. Table 5.1 shows, for both the experimental and the control group, the percentaged means and standard deviations of total 32-item pretest scores, the four logical-form subtest scores (eight items each), and the Stanford Achievement Test total battery scores taken four months prior to the experiment.

Using the variable-by-variable data given in table 5.1, the hypothesis of equal means for the experimental and the control groups is not rejected even at .05 level, for each variable. This is true for both pretest scores and the independent measure of general school work through the Stanford Achievement Test. Based on the assumption that there is no critical significant difference between the two groups initially, the analysis proceeds.

Table 5.1 also shows (in accordance with previous studies cited in section 1.4) that there was indeed room for improvement on participating students' ability to answer AC and DA items, namely, to recognize conclusions that do not necessarily

---

*See section 4.1.1 page 112.

Table 5.1 Experimental vs. Control Groups Pretest and Stanford Achievement Test (S.A.T.) Percentaged Mean Score

Group / Test		Experimental group (n=104)		Control group (n=106)		t - statistics for testing the differences between uncorrelated means*
		Mean (%)	SD	Mean (%)	SD	
Total pretest		54.3	15.0	53.8	13.4	.1
Logical forms subtests	MP	83.3	15.3	81.1	16.1	.2
	MT	78.3	19.1	78.8	18.3	.0
	AC	24.3	30.3	22.1	24.9	.4
	DA	31.3	29.3	33.0	26.9	.3
S.A.T.		69.6	14.7	68.9	14.8	.1

*$t_{208}(.01) = 2.60$; $t_{208}(.05) = 1.97$

follow from the premises. Mean scores on AC and DA subtests were 24.3% and 31.3%, respectively. This is below chance level even if we consider the three alternative answers -- yes/no/not-enough-clues -- as having equal chance.**

Experimental and control groups were further compared on the pretest item-by-item answer profiles. The data is given in table 5.2.

**The assumption of equal chance is on the conservative side. As discussed in section 3.1.2, the answer "not enough clues" (NEC) may raise psychological difficulties in being interpreted as admittance of personal inability to answer the question rather than as inherent logical undecidability of the question. This reluctance to choose NEC as an answer may be increased by the rarity of the occurrence of this possibility in the student's previous school experience. The introductory examples given in the test were intended to help students overcome this psychological difficulty.

Table 5.2

Percentages of Experimental and Control Group Students Selecting Each Answer for the Various Pretest Items.
(Right answers are circled. Errors expected from model in section 1.7.3 are underlined.)

Item No.	Logical Form	Negation Mode	Pretest Frequencies (in percentages)							
			Experimental				Control			
			Yes	No	NEC	Skipped	Yes	No	NEC	Skipped
1	MP	++	88.5	3.8	6.7	1.0	83.1	7.5	9.4	.0
23		++	79.8	12.5	7.7	.0	84.0	11.3	4.7	.0
25		+-	29.8	57.7	12.5	.0	28.3	55.7	16.0	.0
30		+-	2.9	91.3	4.8	1.0	2.9	89.6	7.5	.0
11		-+	91.3	1.0	7.7	.0	88.7	5.7	5.6	.0
13		-+	79.8	8.7	11.5	.0	75.5	13.2	11.3	.0
19		--	1.9	93.3	3.8	1.0	1.9	96.2	1.9	.0
3		--	7.7	84.6	7.7	.0	16.0	76.4	7.6	.0
22	MT	++	1.0	93.3	5.7	.0	3.8	93.3	2.9	.0
31		++	1.9	85.6	9.6	2.9	1.9	84.9	13.2	.0
8		+-	16.3	66.4	15.4	1.9	18.9	61.3	19.8	.0
16		+-	2.9	91.3	5.8	.0	4.7	92.5	2.8	.0
5		-+	51.0	31.7	15.4	1.9	54.7	31.1	14.2	.0
9		-+	85.5	1.0	13.5	.0	87.7	4.7	7.6	.0
20		--	83.7	12.5	3.8	.0	80.2	15.1	4.7	.0
28		--	69.3	25.0	5.7	.0	75.5	20.7	3.8	.0
2	AC	++	76.0	1.0	22.0	1.0	75.5	1.9	22.6	.0
14		++	76.9	4.8	15.4	2.9	70.7	8.5	20.8	.0
4		+-	63.5	1.0	31.7	3.8	69.8	1.0	29.2	.0
7		+-	76.0	1.9	22.1	.0	75.4	5.7	18.9	.0
21		-+	40.4	29.8	29.8	.0	42.4	27.4	30.2	.0
27		-+	1.0	74.0	24.0	1.0	2.8	77.4	19.8	.0
12		--	29.8	45.2	25.0	.0	31.1	50.9	18.0	.0
18		--	2.9	72.1	24.0	1.0	1.9	80.2	17.9	.0
17	DA	++	4.8	68.2	26.0	1.0	6.6	74.5	18.9	.0
10		++	5.8	64.4	28.8	1.0	7.6	74.5	17.9	.0
6		+-	47.1	17.3	35.6	.0	40.6	9.4	50.0	.0
32		+-	28.8	35.6	32.7	2.9	28.3	43.4	28.3	.0
29		-+	3.8	63.5	28.8	3.9	5.7	69.8	24.5	.0
15		-+	7.7	50.0	42.3	.0	8.5	27.4	64.1	.0
24		--	51.9	8.7	39.4	.0	51.9	8.5	39.6	.0
26		--	78.8	4.9	16.3	.0	78.3	.9	20.8	.0

χ^2 test for experimental and control group profiles of identical items showed no significant difference even at $\alpha = .05$ level on all items except item number 15. For this item the hypotheses of equal profiles was not rejected at $\alpha = .025$.

5.1.2 Internal consistency of the test. The pretest consisted of 32 items in 16 item-pairs, identical in terms of logical form and negation mode (see table 3.4). Table 5.3 shows the number of yes/no/not-enough-clues answers given to each item by all 210 participating students.*

The χ^2 test failed to reject the hypothesis of equal profiles for type-mate items only in three of the sixteen pairs, namely in items 1-23, 20-28, 24-26. In all other type mates the discrepancy was significant.** In view of the structure of the test and its previous revisions (see section 3.1.5), inconsistencies between type-mate items could only be attributed to language and content effects as discussed in section 1.4.3 and 1.4.4.

Upon a review of the items themselves (appendix 7.2) some cases of inconsistency could be explained. For other items, no reason for the inconsistencies could be found. The inconsistencies are listed on page 128.

---

*Form T' was identical to form T of the test except for the order of typemates. Therefore they were considered as identical forms for all purposes of analysis. However, by necessity, the difference in order was taken into account in the process of recording the answers. E.g., because item 1 of form T is identical to item 23 of form T', answers to item 23 on T' were recorded as answers to item 1.

**It should be noted that despite the inconsistencies, split-half reliability of the test was .75 and test-retest reliability was .79 (see section 3.2.2).

Table 5.3

Test-Item Profiles Based on Pretest Results of 210 Participating Students.

(Right answers are circled. Errors expected from model in section 1.7.3 are underlined.)

Item No. (in T form)	Logical Form	Negation Mode	No. of Yes's	No. of No's	No. of NEC's	No. Skipped	Total
1	MP	++	180	12	17	1	210
23		++	172	25	13	0	210
25		+-	61	119	30	0	210
30		+-	6	190	13	1	210
11		-+	189	7	14	0	210
13		-+	163	23	24	0	210
19		--	4	199	6	1	210
3		--	25	169	16	0	210
22	MT	++	5	196	9	0	210
31		++	4	179	24	3	210
8		+-	37	134	37	2	210
16		+-	8	193	9	0	210
5		-+	111	66	31	2	210
9		-+	182	6	22	0	210
20		--	172	29	9	0	210
28		--	152	48	10	0	210
2	AC	++	159	3	47	1	210
14		++	155	14	38	3	210
4		+-	140	2	64	4	210
7		+-	159	8	43	0	210
21		-+	87	60	63	0	210
27		-+	4	159	46	1	210
12		--	64	101	45	0	210
18		--	5	160	44	1	210
17	DA	++	12	150	47	1	210
10		++	14	146	49	1	210
6		+-	92	28	90	0	210
32		+-	60	83	64	3	210
29		-+	10	140	56	4	210
15		-+	17	81	112	0	210
24		--	109	18	83	0	210
26		--	165	6	39	0	210

Items	Inconsistency	Inconsistency attributable to
6 and 32*	Item 6 brought the second highest number of right answers in DA. Item 32 does not conform with the predicted error pattern given in section 1.7.3	A conjecture is that underlining of the word "not" in both clues, by atmosphere, may have caused the high number of negative answers. (This was a typographical error because the "not" in the first clue should not have been underlined. See comment on underlined negations in section 3.1.5 page 107).
8* and 16	Item 8 brought the second lowest number of right answers in MT. Item 16 brought the second highest.	Unexplained. Notice: Item 16 includes an implicit universal quantifier. Item 8 does not.
5* and 9*	Item 5 brought the lowest number of right answers in MT. Item 9 brought the third highest number.	Unexplained.
12 and 18*	Inconsistency in wrong answers pattern.	Higher number of wrong yes answers in item 12 as compared to 18 may be attributable to difference in content familiarity in favor of item 18.
15 and 29*	Item 15 brought the highest number of right answers in DA.	Unexplained.
21 and 27*	Inconsistency in wrong answer pattern. Item 21 does not follow the prediction model in section 1.7.3.	Unexplained.
25 and 30*	Item 25 brought the lowest number of right answers in MP. Item 30 brought the third highest number.	Ignoring the conditional sentence and relying only on the second clue may lead to NEC answer on item 25 due to everyday experience. No explanation for the high number of wrong yes answers was found. Notice: Item 30 includes an implicit quantifier, item 25 does not.

---

*This item was revised after the pilot study (see section 3.1.5).

5.2 Posttest Scores

Similar to the data given in table 5.2 for the pretest, table 5.4 gives item-by-item posttest profiles for the experimental and the control groups. Again the numbers are percentaged because the groups were of unequal size (104 students in the experimental group, and 106 in the control group).

χ^2 test shows a significant difference (at $\alpha = .001$) between item profiles of the experimental and the congrol groups on every AC and DA item. In MP and MT the difference is significant at $\alpha = .01$ on items 9,11 and at $\alpha = .05$ on items 5,8,22,30.

Taking into account the great similarity found between the profiles of the two groups on every pretest item, the above results clearly indicate a consistent behavior modification on the part of the experimental group in the two undecidable subtests. These are the subtests the students had great difficulty with in the pretest.

5.3 Pretest/Posttest Comparison, a Descriptive Analysis

5.3.1 Overall comparison. The 32-item pretest and posttest were administered to the experimental and the control groups as described in sections 3.2.3 and 4.3.4. Table 5.5 gives the distribution of overall test scores for the two groups (page 131).

A t-test on posttest mean scores shows that posttest mean scores of the experimental and the control groups differ significantly (at $\alpha = .01$). Pretest to posttest changes in the experimental-group distribution polygon as compared to those changes in the control group polygon are shown in diagram 5.1 (page 132).

Table 5.4

Percentages of Experimental and Control Group Students Per Answer for each Posttest Item.
(Right answers are circled. Errors expected from model in section 1.7.3 are underlined.)

Item No.	Logical Form	Negation Mode	Posttest Frequencies (in percentages)							
			Experimental Group				Control Group			
			Yes	No	NEC	Skipped	Yes	No	NEC	Skipped
1	MP	++	90.4	6.7	2.9	.0	87.7	8.6	2.8	.9
23		++	85.6	2.9	11.5	.0	89.6	2.8	7.6	.0
25		+-	4.8	81.7	12.5	1.0	2.8	91.5	5.7	.0
30		+-	10.6	79.8	9.6	.0	19.8	63.2	16.1	.9
11		-+	86.5	4.8	8.7	.0	67.0	11.3	21.7	.0
13		-+	80.8	1.9	17.3	.0	88.7	2.8	7.6	.9
19		--	5.8	78.8	13.5	1.9	9.4	78.3	12.3	.0
3		--	2.9	83.6	13.5	.0	9.4	84.9	5.7	.0
22	MT	++	1.9	80.8	15.4	1.9	1.9	90.6	7.5	.0
31		++	6.7	84.6	8.7	.0	1.9	88.7	7.5	1.9
8		+-	5.8	78.8	14.4	1.0	9.5	85.8	4.7	.0
16		+-	13.5	73.0	13.5	.0	17.9	76.4	5.7	.0
5		-+	79.8	4.8	14.4	1.0	93.4	1.9	4.7	.0
9		-+	76.0	9.5	13.5	1.0	55.7	30.2	13.2	.9
20		--	73.1	16.3	9.6	1.0	72.6	22.7	4.7	.0
28		--	70.2	15.4	14.4	.0	77.4	15.1	7.5	.0
2	AC	++	34.6	2.9	61.5	1.0	78.3	6.6	15.1	.0
14		++	33.7	.0	65.3	1.0	76.4	4.7	18.9	.0
4		+-	26.0	1.9	70.2	1.9	78.3	1.9	18.9	.9
7		+-	17.3	1.9	80.8	.0	64.1	2.9	32.1	.9
21		-+	3.8	19.2	76.0	1.0	6.6	63.2	30.2	.0
27		-+	9.6	29.8	59.6	1.0	28.3	41.6	29.2	.9
12		--	5.8	26.9	67.3	.0	.9	76.4	22.7	.0
18		--	10.6	26.9	62.5	.0	23.6	52.8	23.6	.0
17	DA	++	4.8	24.0	71.2	.0	1.9	67.0	29.2	1.9
10		++	.0	33.7	66.3	.0	6.6	68.9	24.5	.0
6		+-	13.5	17.3	66.3	2.9	23.6	36.8	38.7	.9
32		+-	19.2	8.7	70.2	1.9	43.4	10.4	44.3	1.9
29		-+	6.7	21.2	71.1	1.0	7.6	35.8	55.7	.9
15		-+	1.9	31.8	66.3	.0	4.7	66.1	28.3	.9
24		--	20.2	5.8	74.0	.0	72.6	3.8	21.7	1.9
26		--	16.3	5.8	76.9	1.0	48.1	4.7	47.2	.0

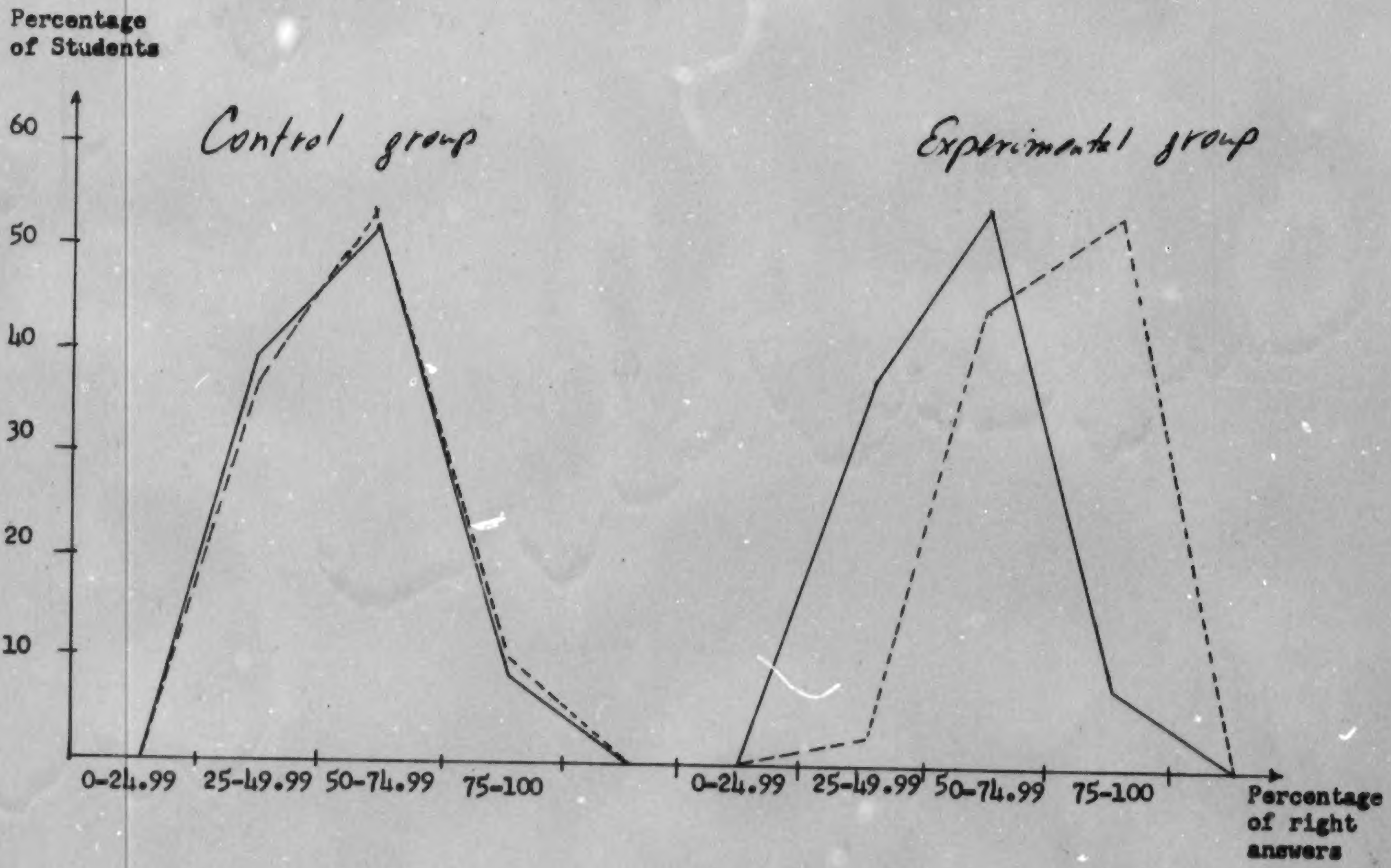
Table 5.5

Distribution (in percentages) of Pretest and Posttest Overall Test Scores for Experimental (n=104) and Control Groups (n=106)

Number of Right Answers (out of 32)	Control (n=106)		Experimental (n=104)	
	Pretest	Posttest	Pretest	Posttest
0 to 7 correct	0	0	0	0
8 to 15 correct	39.7	36.8	37.5	2.9
16 to 23 correct	50.9	52.8	53.8	44.2
24 to 32 correct	9.4	10.4	8.7	52.9
TOTAL	100%	100%	100%	100%
Mean total number correct	53.8%	55.4%	54.3%	74.7%

As diagram 5.1 shows, there was a negligible right-hand shift in the distribution of the control group, i.e., a negligible pretest/posttest improvement on the part of the control group students. Pre- and posttest distributions for this group both have almost identical skewed, bell-shape form, and the maximum frequency in the 50-74.99% set. The experimental group pretest distribution is similar to that of the control group, but its posttest score distribution showed a marked shift to the right with the maximum at the 75-100% right answers set. A decrease of 34.6 in the 25-49.99% right answers with an additional decrease of 9.6 in the 50-74.99% set make an increase of 44.2 in the frequency of the set of 75-100% right answers, jumping from 8.7% on the pretest to 52.9% on the posttest. Significance of the differences needs no statistical establishment. Of more importance

Diagram 5.1 Distribution Polygon of Pretest (———) and Posttest (- - - -) total Score



is the educational significance of these results. Namely 52.9% of the experimental group students consistently gave right answers on the posttest as opposed to 8.7% on the pretest. In other words, 44.2% of the students learned to consistently distinguish valid from non-valid inferences.

5.3.2 Logical form subtests results. The 32-item test was subdivided into four subtests in two ways, thus forming two sets each of four 8-item subtests: Logical Form subtests: (i) MP, (ii) MT, (iii) AC, (iv) DA; and Negation Mode Subtests: (i) ++, (ii) +-, (iii) -+, (iv) --. (For specific items in each subtest see table 3.4.)

Table 5.6 shows pretest and posttest score (number of right answers) distribution for the four logical-form subtests. Frequencies are percentaged in order to equate the unequal sample sizes of the experimental (n=104) and the control (n=106) groups.

Table 5.6

Percentages of Students in the Experimental and Control Groups Having Various Scores on the Logical Form Subtests in the Pretest and the Posttest

Logical Form		MP			MT			AC			DA		
Score Range		0-2	3-5	6-8	0-2	3-5	6-8	0-2	3-5	6-8	0-2	3-5	6-8
PRE	Exp.	1.0	11.5	87.5	2.9	20.2	76.9	73.0	13.5	13.5	57.7	30.8	11.5
	Cont.	.9	20.7	78.4	1.9	20.7	77.4	71.7	21.7	6.6	58.5	27.4	14.1
POST	Exp.	1.9	15.4	82.7	2.9	19.2	77.9	11.5	34.6	53.9	6.7	42.3	51.0
	Cont.	.9	20.7	78.4	.9	18.9	80.2	65.1	27.4	7.5	49.1	34.9	16.0

Diagram 5.2 and 5.3 are pictorial presentations of the data in table 5.6. First, diagram 5.2 (control group) will be examined, then it will be compared with diagram 5.3 (experimental group), and finally diagram 5.3 will be examined independently. Diagram 5.2 shows a great similarity of pretest and posttest graph shapes for the control group on all four logical forms. Nevertheless there is a major difference between the decidable subtests -- MP and MT logical forms -- where the right answer was either yes or no, and the undecidable subtests -- AC and DA logical forms -- where the right answer was: "not enough clues." In the decidable subtests a great majority of the control group students answered 6-8 ($\frac{3}{4}$ or more) items right, 78.4% and 77.4% in the pretest and 78.4% and 80.2% in the posttest, on MP and MT, respectively.

On the other hand, in the undecidable subtests the majority of the control-group students answered only 0-2 ($\frac{1}{4}$ or less) items right, 71.7% and 58.5% of the students in the pretest and 65.1% and 49.1% in the posttest on AC and DA, respectively. There is a negligible pretest/posttest shift to the right in MT and to the left in MP and a slight righthand shift in AC and DA on the part of the control group. This may be attributable to learning from the pretest, because for this group the posttest was a second major confrontation with undecidable problems so rarely encountered in regular school curricula. (These shifts are analyzed quantitatively in section 5.4.2 and their insignificance established statistically.)

A comparison of diagrams 5.2 and 5.3 for pretest distributions of the control and the experimental groups shows a similar starting point for both groups, with a small advantage for the experimental

Diagram 5.2 Control Group Distribution of Pretest (▯) and Posttest (▨) Scores on the Logical Form Subtests

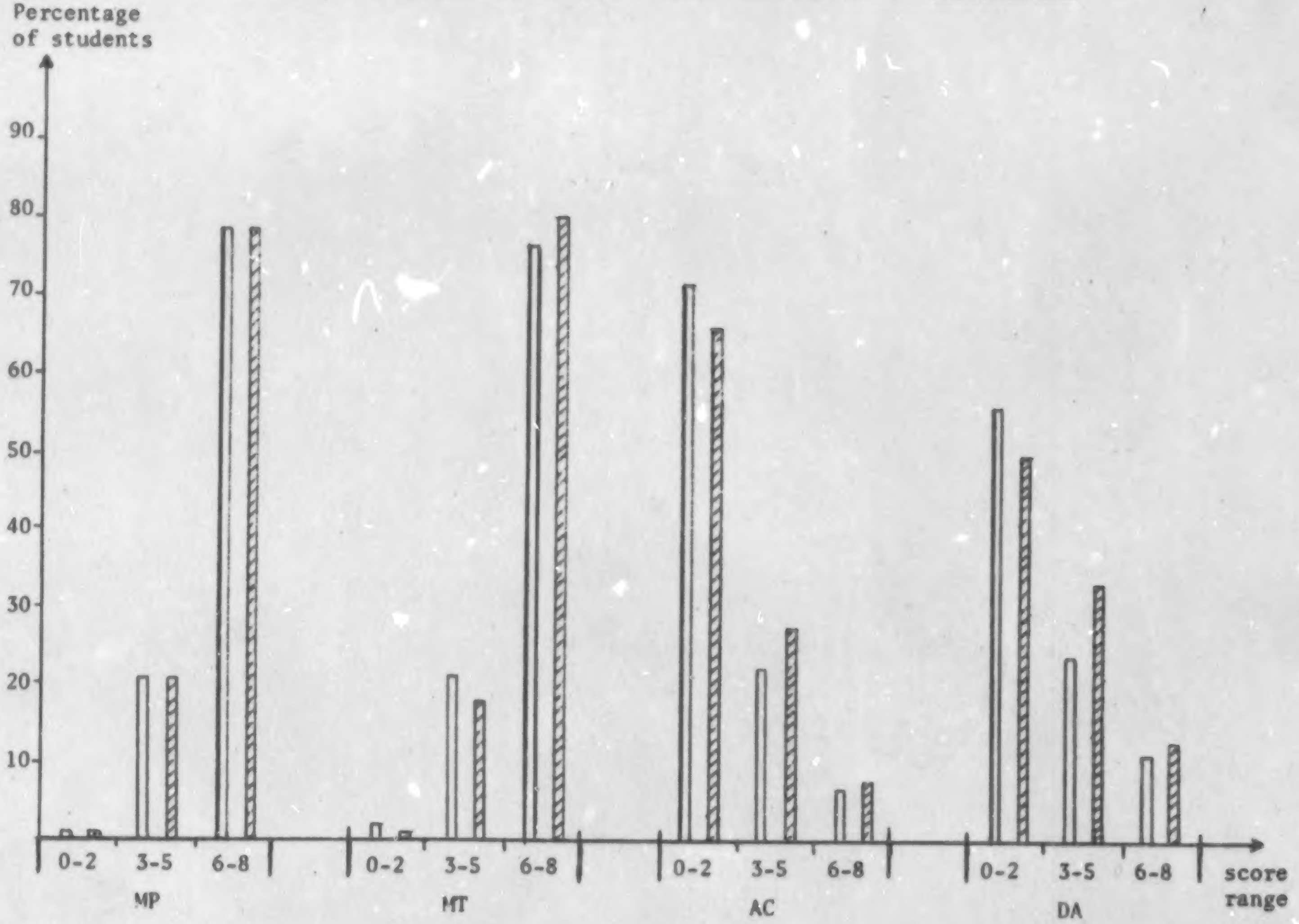
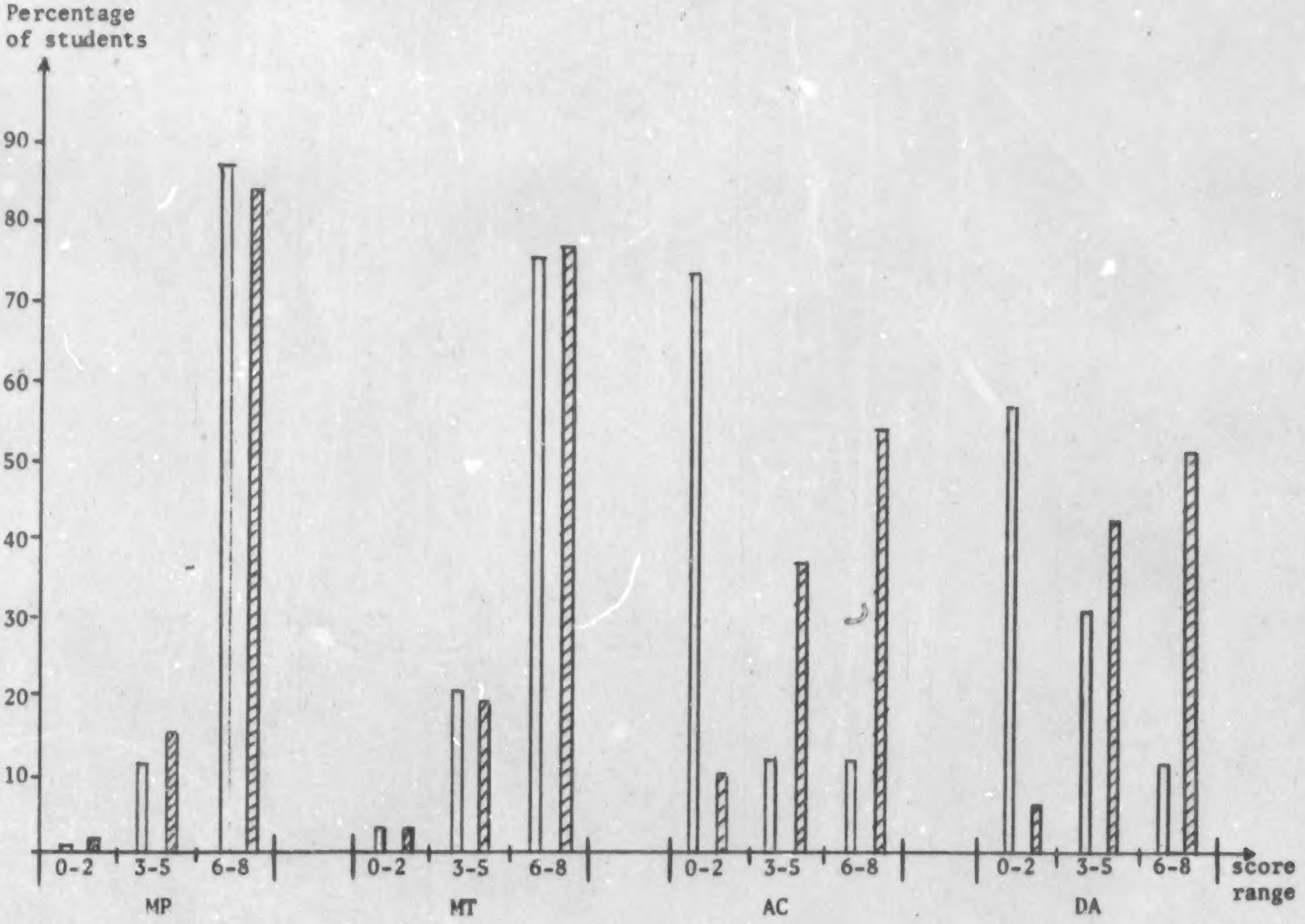


Diagram 5.3 Experimental Group Distribution of Pretest (▯) and Posttest (▨) Scores on the Logical Form Subtests



group on MP, and for the control group on MT, lower extreme scores on AC for the control and, on DA for the experimental group. This reinforces the two groups comparability, previously established by total mean score, which will be further analyzed by subtests in section 5.4.1. However, looking across logical forms in diagram 5.3 for a pretest/posttest comparison of experimental-group students, the tremendous shift in AC and DA right-answer distribution, with the high initial mastery of MP and MT preserved, is quite apparent. A statistical analysis establishing the significance of this shift appears in section 5.4.2. True, posttest mastery level of AC and DA did not reach the mastery level of MP and MT. It is, however, safe to say that major progress was made in the experimental-group students' ability to recognize an unnecessary conclusion from given premises, and to distinguish between necessary and unnecessary conclusions. The educational significance of such an achievement was broadly discussed in Chapter 1.

5.3.3 The individual's progress. So far, only group results have been analyzed. It was found that the experimental group as a whole made significant improvement on the undecidable subtests (AC and DA) and made no significant improvement on the decidable subtests, on which the group performed with relative success on the pretest. Table 5.7 gives the correlation of low, medium, or high performance on pre- and posttests. In each 3 × 3 subtable, the numbers in the diagonal, from the upper left cell to the bottom right cell, are of those students who performed equally well on both tests. Below the diagonal are numbers of students who

Table 5.7 Distribution of Pretest versus Posttest Performance Levels on Each Logical-Form Subtest.

Sub-test	Pretest Level	Experimental Group Posttest Levels				Control Group Posttest Levels			
		low 0-2	average 3-5	high 6-8	total	low 0-2	average 3-5	high 6-8	total
MP	low 0-2	0	1	0	1	0	1	0	1
	average 3-5	0	2	10	12	0	10	12	22
	high 6-8	2	14	75	91	1	11	71	83
	total	2	16	86	104	1	22	83	106
MT	low 0-2	0	1	2	3	0	1	1	2
	average 3-5	2	7	12	21	1	7	14	22
	high 6-8	1	12	67	80	0	12	70	82
	total	3	20	81	104	1	20	85	106
AC	low 0-2	12	30	34	76	63	12	1	76
	average 3-5	0	5	9	14	6	12	5	23
	high 6-8	0	1	13	14	0	5	2	7
	total	12	36	56	104	69	29	8	106
DA	low 0-2	7	31	22	60	44	16	2	62
	average 3-5	0	11	21	32	7	15	7	29
	high 6-8	0	2	10	12	1	6	8	15
	total	7	44	53	104	52	37	17	106

regressed and above it -- those who progressed.

Each subtest will now be considered separately. The discussion begins with MP and MT subtests, but the more interesting findings are in AC and DA subtests.

MP: 75 of the 104 experimental group students (72.1%), and 71 of the control group students (67%) solved 6-8 items right on both the pre- and the posttest. 9.6% experimental students and 11.3% control students improved from 3-5 right answers on the pretest to 6-8 right answers on the posttest. Regression from 6-8 right pretest answers to 3-5 right posttest answers occurred in 13.5% of the experimental group students and 10.4% of the control group students. Bowker's test (Marascuilo, 1976, Chapter 7) for both groups at overall $\alpha =$ .05 did not reject the hypothesis of no pretest to posttest progress.*

MT: The situation in this subtest is very similar to that of MP. 74 experimental students (71.2%) and 77 control students (73.6%) did not change their performance level, of which none had only 0-2 right answers in both tests, seven students in each group had 3-5 right answers in both tests and all the others -- 64.4% of the experimental and 66.0% of the control, had 6-8 right answers on the pretest as well as on the posttest. Here, as in the MP case, Bowker's test for both groups failed at $\alpha = .05$ to reject the hypothesis of no pretest posttest differences.*

---

*One should keep in mind that the no change in the observed results does not necessarily mean that there really was no change in students' understanding. In particular, by Henkin's conjecture (see section 1.7.3), students might get right answers on MP and MT pretest due to a fortunate but a wrong process.

MP and MT: On the whole, MP and MT distributions of control and experimental groups are very similar to each other with respect to individuals' differences. The two independent sample χ^2 tests (Siegel 1956, Chapter 6) on the nine categories tabulated in table 5.7 for each subtest did not reject the hypothesis that the two samples don't fit each other for either MP or MT, even when allowing the confidence level to be very low.** This means that the experimental unit had no significant effect on students' *observed* change in performance on either of the decidable subtests. Such observed changes usually indicate random small fluctuations.**

AC: The situation in this subtest is completely different from the two previous ones. First of all, only one experimental student regressed from the initial performance level. Thirty students (28.8%) remained at pretest performance level; thirteen performed on the high level initially, and twelve (11.5%) remained at the low level of performance with no apparent benefit from the program. However, 73 experimental students (70%) did make progress: 34 (32.7%) jumped from the low to the high performance level and an additional 9 (8.5%) passed to the high level from the average level. Was it due to the effect of the experimental unit? A look at the control group part of table 5.7 for this subtest suggests a positive answer.

---

*See footnote on page 139.

**$\chi^2_8(.10) = 13.36$.

Sixty-three (59.4%) of the control group students started and remained on a low performance level. An additional 14 students (13.2%) at other levels did not change their performance level; altogether 72.6% stayed at their initial performance level (compare to 70% in the experimental group who made progress). One student jumped from the low to the high performance level and the remainder is divided into 17 (16%) and 11 (10.4%), respectively, who made progress or regressed in one category.

Posttest distributions of the experimental and the control groups were found significantly different (at $\alpha = .001$) by the two-independent sample χ^2 test (Siegel 1956, Chapter 6). Pretest to posttest progress on the part of the experimental group was found significant (at overall $\alpha = .05$) by Bowker's test in all three possible progress cells: (1) low pretest level and average posttest level; (2) low pretest level and high posttest level; and (3) average pretest level and high posttest level.

DA: The picture for the DA subtest is very similar to that for the AC subtest. Two experimental students regressed from high to average performance level, 28 (26.9%) remained at the same level, of whom 10 (9.6%) started off on a high level, and only 7 (6.7%) started low and did not benefit from the experimental unit at all. A great majority of the experimental group, 74 students (71.2%), changed their performance level to a higher one; 22 of them (21.1%) went from low to high performance level. Significance of all three possible changes in performance level was established by Bowker's test (Marascuilo 1975, Chapter 7) at overall $\alpha = .05$.

In the control group, on the other hand, most of the students remained at their initial level -- 44 (41.5%) on the low level, 15 (14.1%) on the average level, and 8 (7.5%) on the high level, altogether 63.1%. Two students jumped from low to high performance and one went in the opposite direction. Twenty-three students (21.7%) moved up one level and 13 (12.3%) moved down one level. All the changes on the part of the control group were not found significant ($\alpha = .05$) by Bowker's test.

A comparison of the experimental and the control groups was carried out using the χ^2-test of two independent samples (Siegel 1956, Chapter 6). It shows a significant difference (at $\alpha = .001$) between the distributions of the two groups reported in table 5.7.

AC and DA: A summary of the two undecidable subtests shows with a high degree of confidence ($\alpha = .001$) a significant effect of the experimental unit on students performance on both of these subtests. Table 5.8 summarizes changes in experimental group performance levels for the 16 undecidable items.

As seen in table 5.8, 65 experimental group students, that is 62.5%, had a low pretest score on the undecidable part of the test. Sixty two of those 65 students moved on

Table 5.8 Distribution of Experimental Group Pretest vs. Posttest Performance Levels on the Total of 16 Undecidable Items (AC,DA)

	Posttest levels / Pretest levels	POSTTEST			
		Low 0-4	Average 5-10	High 11-16	Total
PRETEST	Low 0-4	3	32	30	65
	Average 5-10	0	4	20	24
	High 11-16	0	2	13	15
	Total	3	38	63	104

the posttest to a higher level of performance. Moreover, about half of those low pretest achievers jumped to the high posttest performance level.

Altogether 82 experimental students, that is 81.6%, changed their pretest performance level to a higher one, 20 students (i.e., 19.2%) stayed at their initial level, 13 of which achieved the high performance level on the pretest, and only 2 students (1.9%) regressed.

The number of high achievers, that is those who consistently chose "not-enough-clues" as an answer when this answer was right, went from 15 (14.4%) in the pretest to 63 (60.5%) on the posttest.*

---

*Further analysis of the right and wrong "not enough clues" answers is given in section 5.5 page 160.

5.3.4 Negation mode subtest results. Each of the four negation mode subtests contained eight items. Table 5.9 gives pretest and posttest distributions of scores in percentages for both experimental and control groups. A clearer picture is obtained from diagrams 5.4 and 5.5 in which these data are pictured. Again, starting with the examination of the distributions for the control group (diagram 5.4), all four negation-mode subtest distributions are more or less bell shaped, having a peak of between 63 and 81 percent on the 3-5 right answers set. There were only minor differences between the pretest and the posttest. The same is true for the pretest distributions of the experimental group on the four negation-mode subtests, shown in diagram 5.5.

Table 5.9 Percentages of Students in the Experimental (n=104) and Control (n=106) Groups Having Various Scores on Negation Modes Subtests

Negation mode		++			+-			-+			--		
Score range		0-2	3-5	6-8	0-2	3-5	6-8	0-2	3-5	6-8	0-2	3-5	6-8
PRE	Exp.	1.9	79.8	18.3	9.6	70.2	20.2	7.7	74.0	18.3	10.6	70.2	19.2
	Cont.	4.7	81.1	14.2	12.3	67.9	19.8	10.4	63.2	26.4	6.6	77.4	16.0
POST	Exp.	1.0	37.5	61.5	1.0	31.7	67.3	1.0	37.5	61.5	1.9	31.7	66.4
	Cont.	1.9	83.0	15.1	10.4	68.9	20.7	4.7	70.8	24.5	4.7	76.4	18.9

Experimental group posttest distributions of these four subtests too show a great similarity in shape to one another. However, they are all shifted to the right, changing from bell-

Diagram 5.4 Control Group Distribution of Pretest (▯) and Posttest (▨) Scores on the Negation Modes Subtests

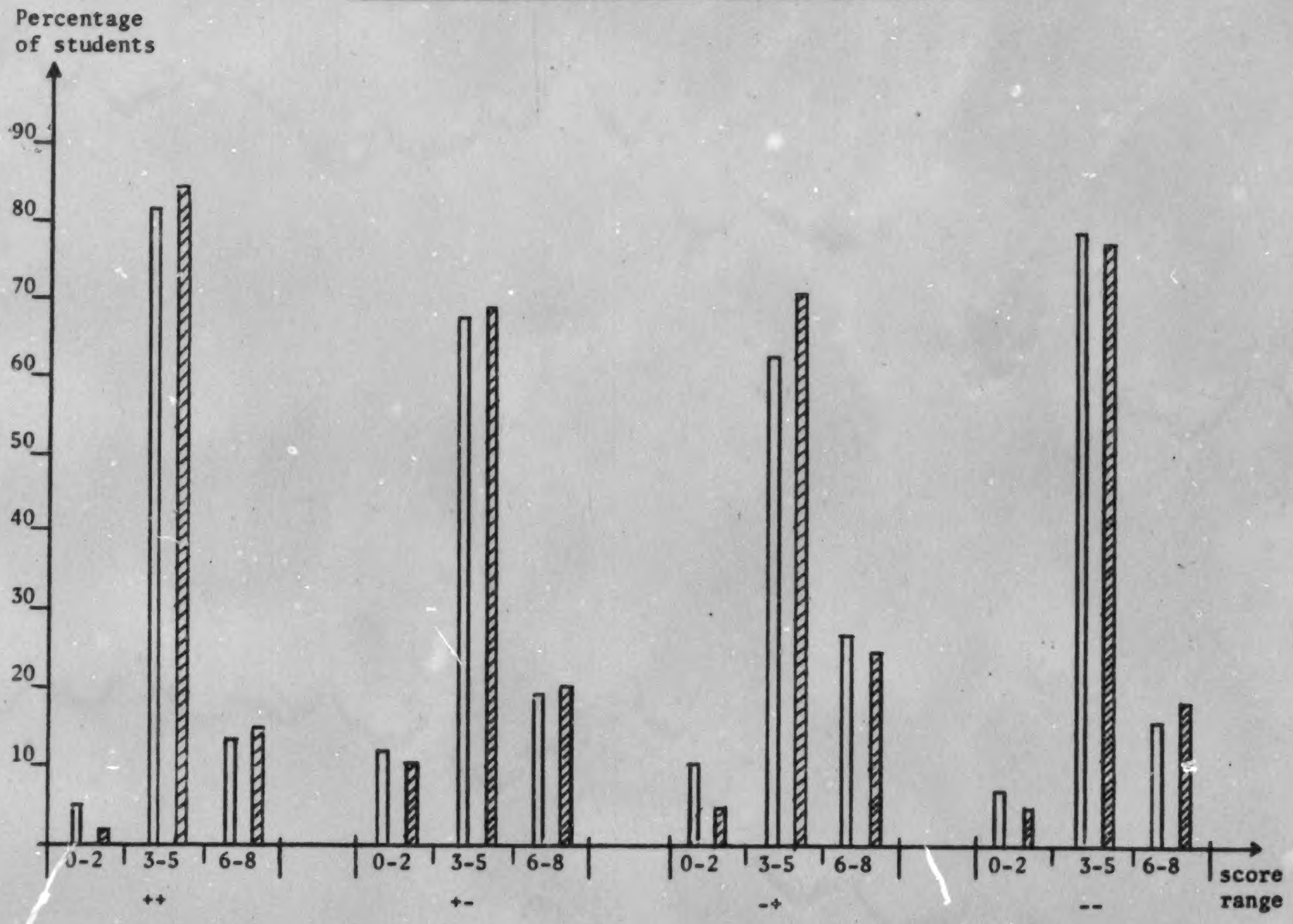
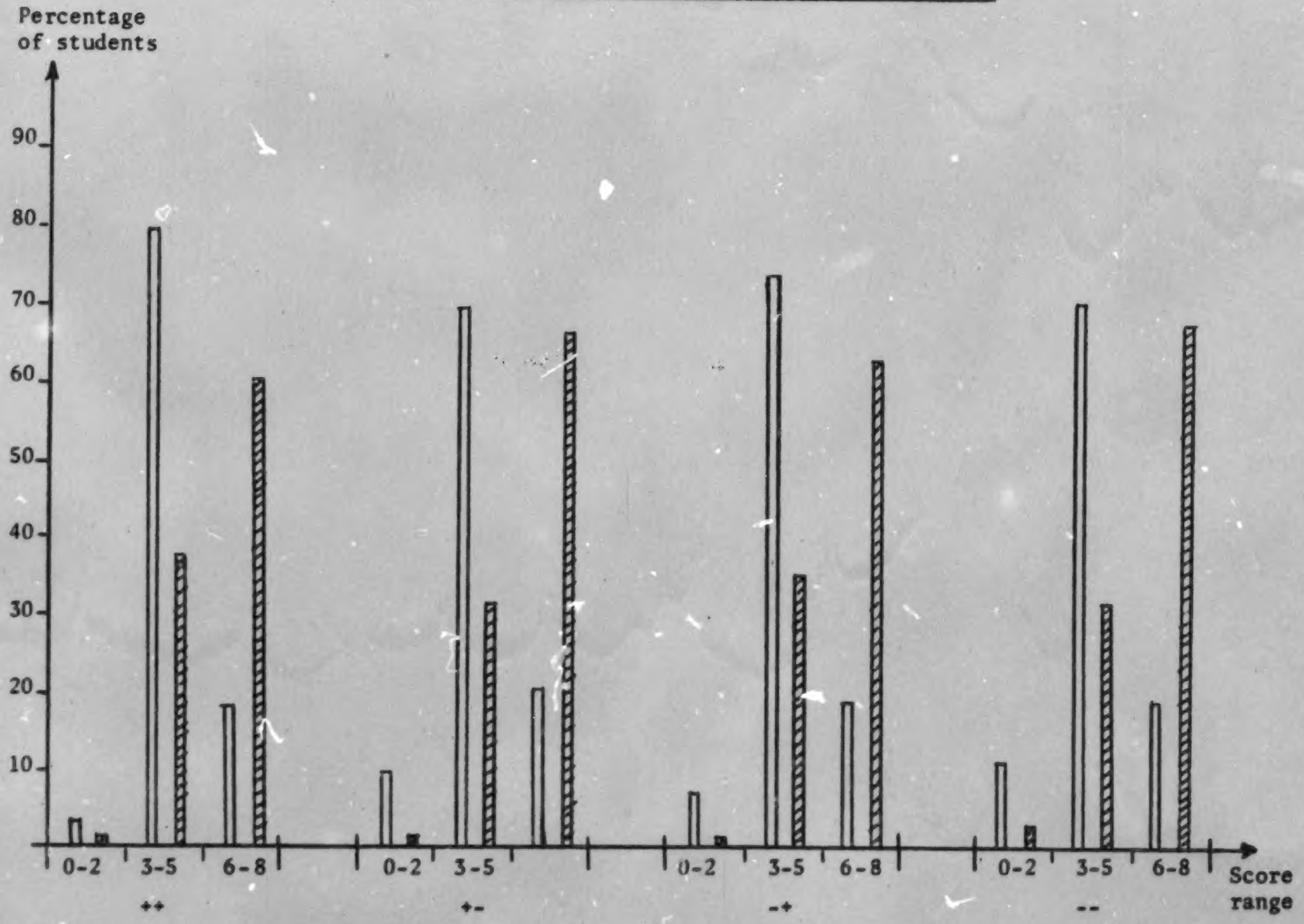


Diagram 5.5 Experimental Group Distribution of Pretest (▯) and Posttest (▨) Scores on the Negation Modes Subtests



shaped functions in the pretest to increasing functions in the posttest. The peak in 3-5 right answers in each negation mode on the pretest distribution decreases, and is roughly halved on the posttest. Zero to 2 pretest frequency diminishes in the posttest. On the other hand, the low 6-8 right answers pretest frequency increases, roughly multiplied by 3 in the posttest. The marked difference is apparent between the experimental group's performance on the posttest and both the experimental group's pretest and the control group's posttest performance. These results are further analyzed in the next section.

5.3.5 A tempting false conclusion about the effect of negation on conditional reasoning. Negation mode was designated to items by the number of times negation occurred in the first premise. Thus negation mode is a language component in the structure of the relevant conditional sentence. The data given in Table 5.9 as pictured in diagrams 5.4 and 5.5 show very small differences among the distributions of the different negation mode scores on pretests and posttests for the experimental as well as for the control group. Recall that negation was found to add special difficulty to logical reasoning (see section 1.4.7). Therefore, the results as presented above are very surprising. It is particularly surprising because the structure of the negation-mode subtests seems to control for the logical form of the items by the fact that each negation mode subtest is composed of four pairs of items, one pair in MP, one in MT, one in AC, and one in DA logical forms. Due to this identity in logical forms composition of the four different negation mode subtests, it is tempting to conclude from the distribution's similarity that negation in the first premise did not make much difference in participating students' performance on the conditional reasoning test as a whole. Moreover, it may even suggest that conditional reasoning is as easy from "++" conditional premise as it is from a "--" one. However, the last conclusion is false. The situation as described is an excellent example of an observed similarity caused by masked differences and not by a true similarity. This phenomenon will be discussed here following the logic behind M. G. Kendall's coefficient of concordance.

Suppose four raters are asked to rate four ratees. Call the raters A,B,C,D, and the ratees W,X,Y,Z. Final judgment on the ratees is based upon the total score given to each ratee by the raters. Consider two extreme cases. (i) All four raters agree on the rank order of all four ratees. (ii) There is not even one ratee on whose rank order any two raters agree. Table 5.10 gives a particular rank order to exemplify each of the two cases.

Table 5.10 Total Agreement and Complete Disagreement Among Hypothetical Raters on Rank Order of Hypothetical Ratees.

Case (i): Total agreement						Case (ii): Complete disagreement					
Ratees	Raters					Ratees	Raters				
	A	B	C	D	Total		A	B	C	D	Total
W	1	1	1	1	4	W	1	4	3	2	10
X	2	2	2	2	8	X	2	1	4	3	10
Y	3	3	3	3	12	Y	3	2	1	4	10
Z	4	4	4	4	16	Z	4	3	2	1	10

The total column in case (i) of table 5.10 shows a great diversity of the total score in case of total agreement among the raters. The total column in case (ii) shows identical total score in the case of complete disagreement among the raters. This is in fact a very logical outcome since in the case of total agreement it could be very easy to grant first prize to the ratee who was judged to be the best one by all four judges. But in the case of total disagreement among the raters it would be extremely difficult to decide who deserves the first prize. The example above illustrates the fact that when raters have exactly equal judgments total scores are extremely different, on a 10 ± 6 range. And when raters have no agreement in judgments all total scores are 10.

Let us come back to the two different partitions of the test used in the present study, into four subtests -- one by negation mode and one by logical form. Mean scores for type-mate pairs of items were computed for the experimental group pretest and posttest separately. Table 5.11 shows the decreasing order of mean scores which may be interpreted as increasing order of difficulty of (i) logical forms within each negation mode, and (ii) negation modes within each logical form. Case (i) of Tables 5.10 and 5.11 are similar to each other. The same holds for case (ii) in these two tables. This similarity brings the discussion back to the tempting conclusion of negligible differences among negation modes. The right hand side of Table 5.11 suggests a different analysis of the results. Total rank order scores in case (ii) of this table is $10\frac{1}{4} \pm \frac{3}{4}$ in the pretest and $9\frac{3}{4} \pm 1\frac{3}{4}$ in the posttest. (Compare with 10 ± 0 in the extreme case of table 5.10.) As explained

Table 5.11 Rank Order of Increasing Difficulty of
(i) Logical Forms Within Negation Modes
(ii) Negation Modes Within Logical Forms

(i) Logical forms rated by negation modes							(ii) Negation modes rated by logical forms						
Ratees		Raters					Ratees		Raters				
		++	+-	-+	--	Total			MP	MT	AC	DA	Total
PRETEST	MP	2	2	1	1	6	PRETEST	++	2	1	4	3	10
	MT	1	1	2	2	6		+-	4	2	1.5	2	9.5
	AC	4	4	4	4	16		-+	3	4	1.5	1	9.5
	DA	3	3	3	3	12		--	1	3	3	4	11
POSTTEST	MP	1	1	1	1	4	POSTTEST	++	1	1	4	2	8
	MT	2	2	2	2	8		+-	4	3	1	3.5	11.5
	AC	4	3	3.5	4	14.5		-+	2	2	2	3.5	9.5
	DA	3	4	3.5	3	13.5		--	3	4	3	1	11

for case (ii) of Table 5.10, the small variance of total scores indicates a great diversity. In the pretest part of Table 5.11 case (ii), there is only one place where two logical forms agree on the rate of any negation mode -- in both MT and AC "--" is third in order of difficulty. DA and MP, however, take two extreme ratings on this negation mode. On the other hand in rating "++", MT and AC take the extreme ratings 1 and 4 respectively, and DA and MP are closer to each other. Adding the ranks across logical forms would cover these opposite trends. To take another example: examine the posttest ranks for "--" and "+-" in the different logical forms. "+-" was found the most successful mode in AC subtest and was the least successful mode in MP subtest, whereas "--" was the third in both of these subtests, was the most successful one in DA, and was the least successful mode in the MT subtest. Despite the fact that rank profiles for "+-" and for "--" are almost at opposite extremes, when their ranks across logical forms are added, the totals appear to be very close: 11.5 for "+-", and 11 for "--". Again, it is the big difference that causes the great similarity. In short, negation indeed did make a difference in subjects' learning of conditional reasoning as measured in the present study.

Turn now to case (i) of tables 5.10 and 5.11. Logical forms are similarly rated by negation modes. This can be seen by comparing ranges of the total columns in the pretest and in the posttest parts of case (i) table 5.11 -- 11 ± 5 and 9.5 ± 5, respectively. (Compare to the ideal case of table 5.10 where the range is 10 ± 6.) Indeed in both the pretest and the posttest parts of case (i) in table 5.11, MP and MT are scored 1 or 2 in each negation mode and AC and DA are scored 3 or 4 in each negation mode. In other words, in each negation mode separately and in both pre- and posttests, decidable problems (MP, MT) were more successfully answered than undecidable ones (AC, DA). This conforms with the results for logical forms subtests added across negation mode, given earlier in section 5.3.2, page 133.

Following the analysis of case (i) in table 5.11, the results from here on will be considered in terms of logical forms across negation modes. Analysis of negation modes across logical forms will be omitted to avoid distorted conclusions.

5.4 Statistical Analysis of the Results

5.4.1 Analysis of variance on pretest scores. Eight fifth-grade classes participated in the main study -- four experimental classes and four control classes. Altogether 210 students were involved, 104 in four experimental classes and 106 in four control classes. An analysis of variance on the pretest scores was designed to determine whether or not the eight classes had a comparable initial performance level in conditional reasoning as measured by the test. Table 5.12 gives the results of this analysis for each logical-form subtest.

Table 5.12 ANOVA on Pretest Scores

Logical type	Source	df	SS	MS	$F_{7,209}(.99)^* = 2.78$ $F_{7,209}(.90)^* = 1.75$
MP	Between classes	7	16.6	2.4	1.6
	Within classes	202	311.6	1.5	
	Total	209			
MT	Between classes	7	14.4	2.1	<1
	Within classes	202	450.96	2.2	
	Total	209			
AC	Between classes	7	43.3	6.2	1.3
	Within classes	202	976.3	4.8	
	Total	209			
	Between classes	7	18.5	2.7	<1
	Within classes	202	1030.1	5.7	
	Total	209			

*F values obtained by interpolation.

According to the results presented in Table 5.12, the hypotheses of equality among the eight classes' means is not rejected for any logical form, even when a higher risk of .10 is taken for type-1 error. It should be noted here that one of the assumptions upon which the analysis of variance rests is not satisfied for the data analyzed above. Namely, scores of each class are not distributed normally. Class distributions are rather similar to whole group distributions given in diagrams 5.2 and 5.3. However, Norton, cited by Guilford and Fruchter (1973), varied the shape of distribution in various ways,

> "making it leptokurtic, rectangular, markedly skewed and even J-shaped...., One general finding was that F is rather insensitive to variations in shape of population distribution."

Table 5.13 gives class and group means and standard deviations for pretest, posttest, and gain scores. As can be seen from this table, in AC the standard deviation is exceptionally low for class 1 of the experimental group. This is the only case where the assumption of equality of variances within classes, which is necessary for execution of analysis of variance, is violated. However, since the means were found to be insignificantly different, this violation should not be important. Only when means are found to be significantly different will violation of that assumption interfere because the difference found may then indeed be attributed to the differences between the variances and not to differences between the means. The above quote of Norton's results continues to say:

Table 5.13 Class Means and Standard Deviations for Pretest, Posttest, and Gain Scores

		EXPERIMENTAL GROUP										CONTROL GROUP									
		MP*		MT*		AC*		DA*		Total**		MP*		MT*		AC*		DA*		Total**	
		Mean	SD	Mean	SD	Mean	SD	Mean	SD	Mean	SD	Mean	SD	Mean	SD	Mean	SD	Mean	SD	Mean	SD
class 1		n=26										n=26									
	Pre	7.0	.9	6.6	1.4	1.2	1.1	2.2	1.7	17.0	3.4	6.5	.9	6.2	1.4	1.9	1.8	2.2	1.8	16.8	3.9
	Post	7.2	1.0	6.3	1.3	4.7	1.9	5.7	1.7	23.9	3.5	6.5	1.5	6.1	1.4	2.0	2.1	2.6	1.9	17.2	4.2
	Gain	.2	1.4	-.3	1.9	3.5	1.8	3.5	1.8	6.9	3.2	.0	1.8	-.1	1.7	.1	2.0	.4	.9	.5	2.2
class 2		n=25										n=24									
	Pre	6.9	.9	6.4	1.3	1.6	2.5	2.3	2.5	17.3	4.7	6.3	1.4	6.4	1.7	1.7	2.0	3.0	2.1	17.4	5.1
	Post	6.7	1.2	6.2	1.3	5.3	2.1	5.6	1.9	23.9	4.2	6.3	1.4	6.5	1.2	1.8	1.9	3.3	2.3	17.9	4.1
	Gain	-.2	1.4	-.2	1.8	3.7	2.8	3.3	2.7	6.6	4.3	.0	1.8	.1	1.6	.1	1.9	.3	2.0	.5	5.2
class 3		n=28										n=25									
	Pre	6.1	1.0	5.8	1.2	2.5	2.9	2.9	1.8	17.3	5.5	6.6	1.5	6.3	1.3	2.2	1.7	2.6	2.2	17.7	3.5
	Post	6.1	1.4	6.1	1.7	5.8	2.0	5.9	1.9	23.8	4.6	6.7	1.3	6.4	1.3	2.4	2.4	2.7	2.6	18.2	3.8
	Gain	.0	1.8	.3	1.6	3.2	2.5	3.0	1.4	6.5	4.1	.1	1.1	.2	1.3	.2	1.9	.1	2.5	.5	2.8
class 4		n=25										n=28									
	Pre	6.7	1.6	6.2	1.9	2.4	2.5	2.6	2.9	17.9	5.2	6.5	1.1	6.4	1.3	1.4	2.3	2.8	2.3	17.0	4.4
	Post	6.8	1.9	6.0	1.1	6.0	2.1	5.2	2.3	24.0	5.3	6.5	1.3	6.6	.9	1.5	2.5	3.0	2.4	17.5	4.4
	Gain	.1	2.2	-.2	1.7	3.6	2.5	2.7	2.7	6.1	5.5	.0	1.5	.2	1.3	.1	1.7	.2	1.6	.5	3.5
Total group		n=104										n=106									
	Pre	6.7		6.3		1.9		2.5		17.4		6.5		6.3		1.8		2.6		17.2	
	Post	6.7		6.2		5.4		5.6		23.9		6.5		6.4		1.9		2.9		17.7	
	Gain	.0		-.1		3.5		3.1		6.5		.0		.1		.1		.3		.5	

*Maximum score = 8 **Maximum score = 32

> "Even if some pair (of variances) shows a significant difference, one may proceed with analysis of variance, but should then discount significance level somewhat. If F proves to be significant at the .05 level, this result may actually indicate significance at levels .04 to .07."

In the above case, differences were not found to be significant even at .10 level, therefore the conclusions are justified, that all classes of the main study were not significantly different initially.

5.4.2 Analysis of variance on gain scores. Based on the findings in the previous section, one may proceed to analyze the variances of posttest scores without considering pretest scores as covariants. However, distributions of posttest scores were not normal (see diagrams 5.2, 5.3). Even though, as mentioned earlier, the analysis of variance is tolerant of violation of the normality assumption, gain scores seemed to provide more adequate data for the following reasons: a) Gain scores are alternative data for measuring the progress students made between tests, b) Gain score distributions were bell-shaped and closer to normal than posttest score distributions as can be seen from diagrams 5.6 and 5.7.

Examination of diagrams 5.6 and 5.7 show that gain-score distributions were centered around zero in all logical forms for the control group and in MP and MT in the experimental group. However in AC and DA for the experimental group the gain is centered between 3 and 4. These results are consistent with those mentioned in section 5.3.2.

Table 5.14 gives the analysis of variance for testing equality of mean gain-scores between experimental and control groups on each logical form subtest.

Diagram 5.6 Experimental Group Distribution of Gain Scores for the Four Logical Form Subtests

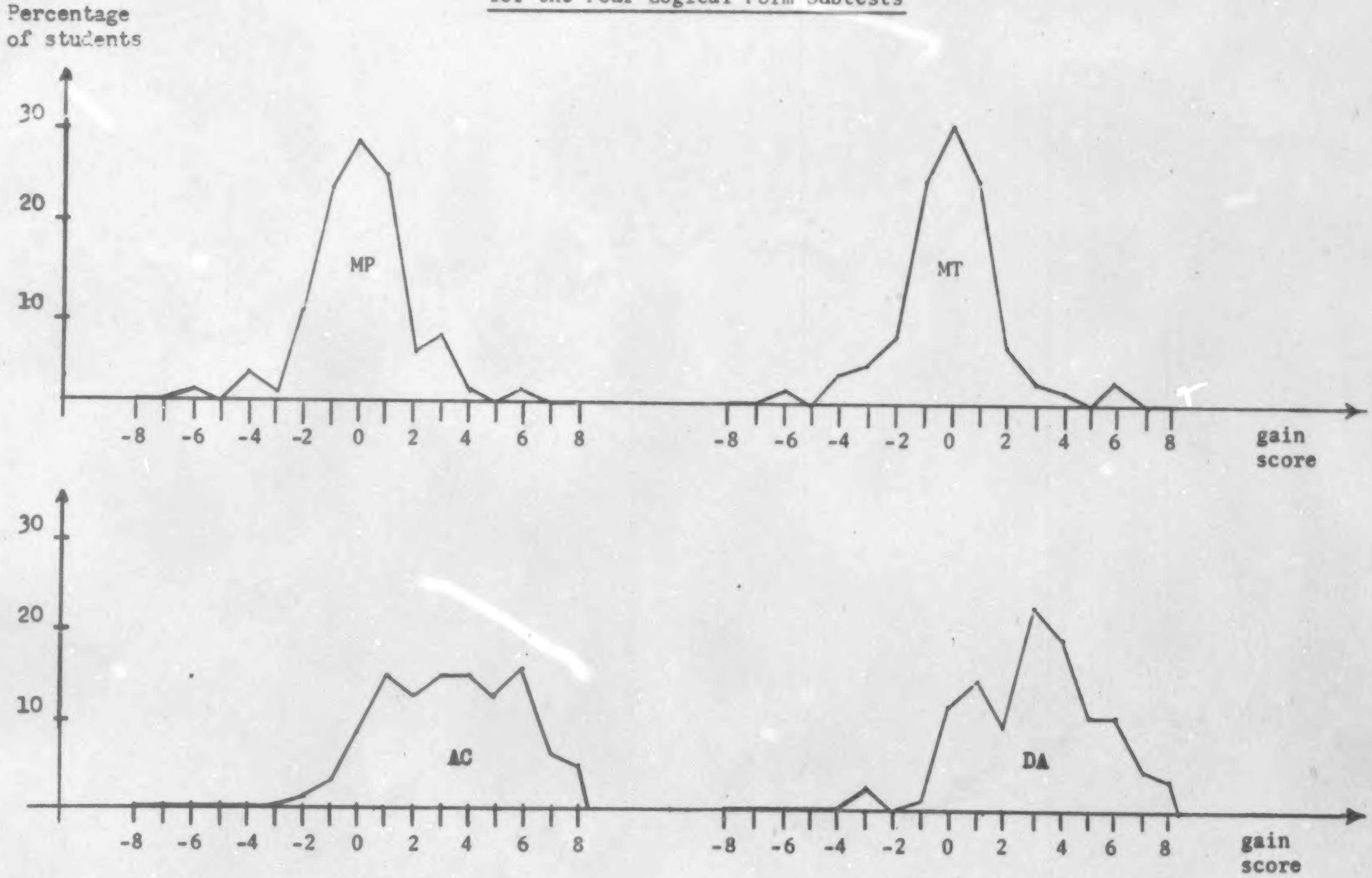


Diagram 5.7 Control Group Distribution of Gain Scores for the Four Logical Form Subtests

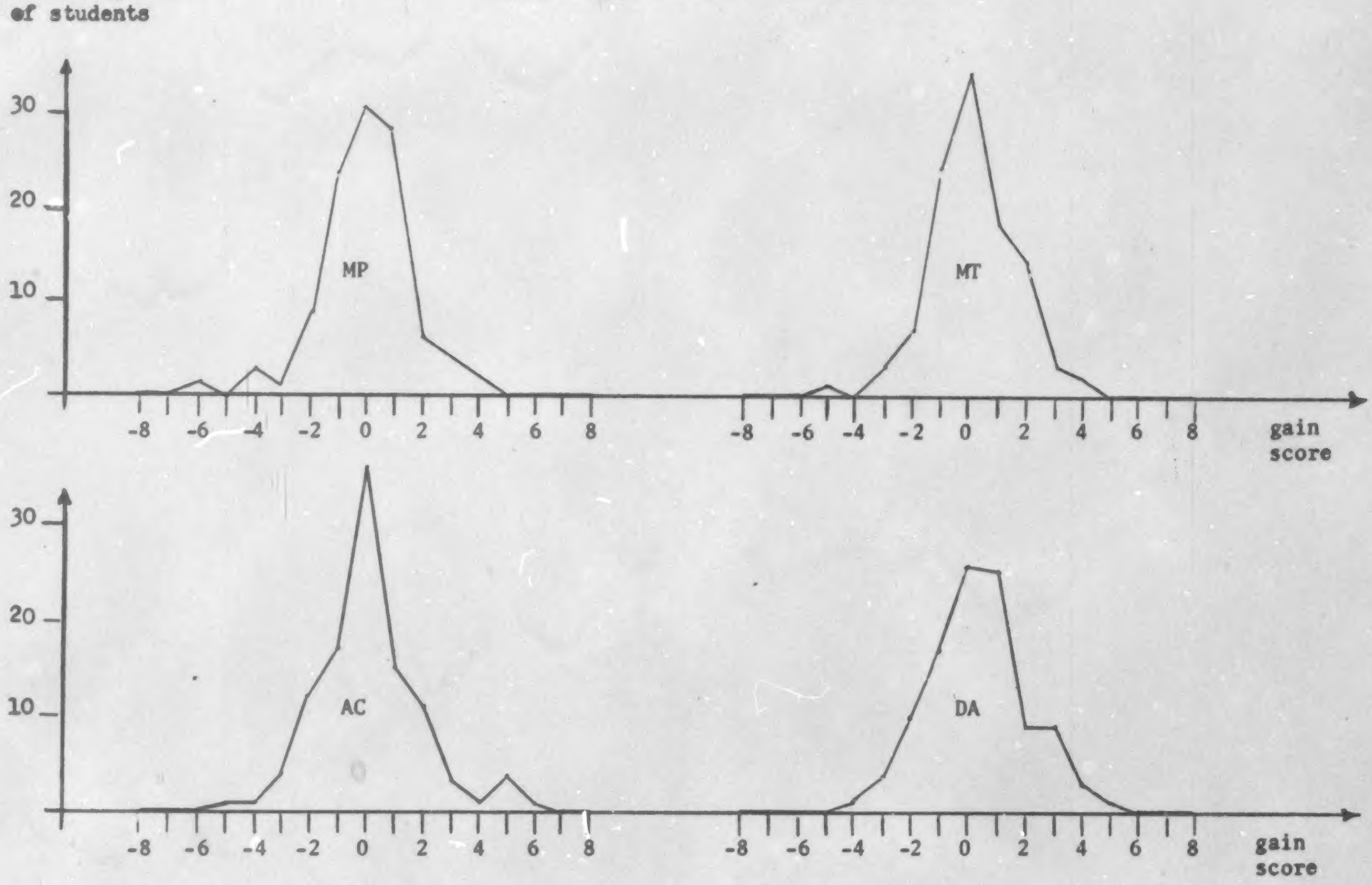


Table 5.14 Analysis of Variance on Pretest to Posttest Gain Scores

Source	df	MP			MT			AC			DA		
		SS	MS	F	SS	MS	F	SS	MS	F	SS	MS	F
Between treatment groups (experimental and control)	1	.01	.01	<1	2.61	2.61	<1	592.65	592.65	125.03*	473.33	473.33	109.06*
Within treatment groups	208	570.9	2.74		560.38	2.69		984.4	4.73		901.51	4.33	
Between classes in experimental group	3	2.04	.68	<1	6.02	2.00	<1	3.76	1.25	<1	9.38	3.12	<1
Within classes in experimental group	100	312.44	3.12		318.2	3.18		602.4	6.02		526.31	5.26	
Between classes in control group	3	.66	.22	<1	1.57	.52	<1	.25	.08	<1	1.29	.43	<1
Within classes in control group	102	255.84	2.51		234.59	2.3		377.99	3.7		364.53	3.57	
Total	209												

*$F_{1,208}(.990) = 6.85$

$F_{3,100}(.990) = 4.48$

$F_{3,102}(.990) = 4.53$

The results of the analysis given in Table 5.14 are as follows:

(a) No significant difference was found between experimental- and control-group mean-gain scores on either MP or MT. Note: pretest scores on these subtests showed a high initial performance level.

(b) No significant difference was found among mean-gain scores of experimental group classes, nor among those of control group classes.

(c) A significant difference was found between the mean-gain scores of the two groups on both AC and DA. However, the hypothesis of equality of variances of the two groups was rejected at .05 level of significance. As mentioned earlier (see section 5.4.1) Norton found the analysis-of-variance tolerant of certain violations of the equality-of-variance assumption, and indicated the need to consider the significance level lower in such a case. In both AC and DA cases, F-ratio is greater than the critical value not only at .01 significance level but even at .001. So, lowering the significance level will still keep the difference significant at .01 level.

This statistical treatment establishes the conclusions one would have drawn intuitively from the data presented in diagrams 5.6 and 5.7. It also establishes the conclusions qualitatively drawn in section 5.1, 5.2, and 5.3.

5.5 Did Students Learn Anything Beyond the Legitimacy of "Not-Enough-Clues" as an Answer?

The psychological difficulties involved in giving the answer "not-enough-clues" were discussed earlier (see footnote on page 124). Throughout the teaching period of the study, experimental-group students were exposed to a variety of problems for which this answer was right. It would not be surprising, therefore, if they had overcome the psychological block due to their new experience in which "not-enough-clues" was legitimized. The purpose of this project was far too ambitious to be satisfied with only this achievement. Therefore, the data was examined in various ways to find indications of learning the logical meaning of "not-enough-clues" which expresses an inherent logical undecidability, over and beyond its legitimacy as an answer.

5.5.1 Analysis of NEC uses by individual students.

The question of right and wront NEC answers on the posttest as compared with that of the pretest for the experimental class was first attacked through individuals' frequencies of choosing it.

For each experimental-group student the number of times he/she answered NEC rightly and wrongly was counted on both the pretest and the posttest. Individuals' pretest to posttest gains were computed for the right uses and for the wrong uses. As 16 items had NEC for the right (wrong) answer, the gain in the right (wrong) NEC uses could go from -16 to +16. Because the desired outcome was a decrease in the number of wrong NEC answers (i.e., a negative wrong use gain), an increase in that number, that is a positive wrong use gain, would be interpreted as overlearning* of the NEC case. To extract the overlearning effect, the difference between NEC right-use gain and wrong-use gain for each student was computed. This difference would be interpreted as an indication of learning of the true meaning of the NEC answer. Distribution of this difference is given in table 5.15.

---

*Overlearning was the term used by teachers for students who learned to increase NEC usage without distinguishing when it was correct.

Table 5.15 Experimental Group Distribution of the Difference between Individual Gains on Right Uses of NEC and Wrong Uses of NEC.

Difference	-32 to -2	-1	0	1	2	3	4	5	6	7	8
Frequency	0	1	0	1	0	3	10	8	17	9	6

Difference	9	10	11	12	13	14	15	16 to 32	omitted
Frequency	5	11	5	3	2	1	2	0	20

Two extreme values of the difference between the right and the wrong use gains are discussed below.

The largest possible value for the difference reported in table 5.15 is 32, which results from a +16 gain in right use of NEC and a -16 gain in wrong NEC use. This extreme value would be interpreted as pure between-tests learning of the undecidability expressed by NEC answer. This is because the only way a student could obtain these gains would be to have on the pretest sixteen (the maximum) wrong NEC and none on the posttest in addition to zero right NEC answers, i.e., no correct AC or DA item correct on the pretest and all sixteen AC and DA items correct on the posttest.

The smallest possible value for the difference reported in table 5.15 is -32 which results from a -16 right use gain and a +16 wrong use gain. Such an outcome is the least desirable and would be interpreted as pure overlearning. Twenty students whose gain in both right and wrong NEC uses were within the ±3 range were omitted from the computation of difference between the two gains because such gains in both right and wrong uses of NEC were considered insignificant in view of the similar fluctuation in the control group.*

Therefore, in general a positive difference in table 5.15 indicates true learning and the larger the positive difference, the more true learning occurred. A negative difference in table 5.15 indicates overlearning of the NEC case and the smaller the figure the more overlearning occurred. Only one student had a negative difference, and this was -1, the minimal overlearning possible. (This student had 4 and 5 right and wrong gains respectively.)

A positive number in table 5.15 was obtained in three cases:

(i) When the right-use and wrong-use gains were positive but the right-use gain was greater than the wrong-use gain. (At least one gain had to be greater than or equal to four.) This was the case in 65 of the 83 positive differences.

(ii) When the right-use gain was positive and the wrong-use gain was negative. (Absolute value of at least one gain was ≥4.) This was the case in 16 of the 83 positive differences.

(iii) When both the right and the wrong gains were negative but the right gain was greater (smaller absolute value). This was the case in 2 of the 83 positive differences; one had a difference of 1 and the other had a difference of 11.

---

*Of these twenty students, eighteen had a right-uses gain greater than or equal to the wrong uses gain. Also, eighteen of these twenty students had a positive right uses gain.

So, 81 experimental-group students, i.e., 77.9%, exhibited learning of the true meaning of NEC to some extent. Sixty of these 81 students, i.e., 58.7% of the experimental group, obtained a difference of 6 or more. It seems safe to infer from the data in table 5.15 that at least 55% of the students learned the logic behind the NEC answer and not just the legitimacy of NEC as an answer.

5.5.2 Analysis of average use of NEC per test item. Using the NEC columns of tables 5.2 and 5.4 (pages 125 and 129), the average number of uses of NEC as an answer per item within each logical form were computed. They are given in table 5.16.

It should be noted that in MP and MT (decidable) subtests, NEC was never the right answer. In all items in the AC and DA (the undecidable) subtests, NEC was the (only) right answer.

Based upon control-group averages, as shown in table 5.16,

Table 5.16 Percent of Students Giving the Answer NEC Per Item (Averaged) in Each Logical Form Subtest.

Logical form	Experimental Group		Control Group	
	Pretest	Posttest	Pretest	Posttest
MP (NEC is wrong)	7.8	9.9	8.0	9.9
MT (NEC is wrong)	9.4	12.9	8.6	6.9
AC (NEC is right)	24.3	67.9	22.1	23.8
DA (NEC is right)	31.3	70.3	33.0	36.2

pretest-posttest fluctuations of ±1.9% would be considered as random fluctuations, resulting from learning from the pretest, differences in weather, the passing of a month's time, etc., etc.

The fluctuations in the number of (wrong) uses of NEC as an answer in the decidable subtests -- MP and MT -- on the part of the experimental group, are of no more than +3.5. This increase in the number of wrong NEC answers would be attributed to overlearning of the NEC case. In fact, because an increase of about 1.9 might be attributed to other random factors that influenced both the control and the experimental groups, only about 1.6 should be considered as resulting from guessing due to legitimization of NEC as an answer. To stay on the conservative side, the total 3.5% will be considered as attributed purely to this factor.

The experimenter considered the number 3.5, interpreted as overlearning of NEC, a low percentage. To separate the learning of the true meaning of NEC as an answer, 3.5 was subtracted from posttest mean scores for the two logical forms in which this answer was correct -- AC and DA. For the experimental group this subtraction still leaves a posttest minus pretest difference of about 40% on AC and of about 35.5% on DA. So, between the pretest and the posttest, an average of 35-40% of the experimental group students per undecidable item learned to choose NEC as an answer indicating that a conclusion does not necessarily follow from the premises.*

---

*Notice that this is an item average. It does not indicate consistency over items. Consistent answers were discussed in sections 5.3.1 for the test as a whole and in section 5.3.2, 5.3.3 for the undecidable items.

5.5.3 Analysis of total NEC uses. Table 5.17 gives the total number of times NEC answers were given rightly and wrongly by all experimental group students in the decidable subtests (MP and MT) and in the undecidable subtests (AC and DA).

There were altogether 16 items in AC and DA subtests. If all items were answered rightly by all 104 experimental group subjects, the number of right uses of "Not-enough-clues" would come to 1664. As table 5.17 shows, the observed number of right uses in the pretest was 462, i.e., 27.8% of the total possible number, whereas in the posttest the total reached 1150, i.e., 69.1%, an increase by the factor of 2.5. If this increase was only due to learning that NEC is a legitimate answer and not to the learning of its logical meaning, the same rate of increase by 2.5 should have occurred also in the wrong uses of this answer. However, as table 5.17 shows, the number of wrong uses of NEC went up only by a

Table 5.17 Observed Number of Right and Wrong Uses of the Answer: Not-Enough-Clues, by Experimental Group Students (n=104)

	MP and MT (16 items)		AC and DA (16 items)		Whole Test (32 items)	
	Pretest	Posttest	Pretest	Posttest	Pretest	Posttest
Right	0*	0*	462	1150	462	1150
Wrong	143	201	0**	0**	143	201
Total	143	201	462	1150	604	1351

*NEC Could not be given rightly in this part of the test (all right answers in this part were yes/no).

**NEC could not be given wrongly in this part of the test (all questions in this part had NEC as their right answer).

factor of 1.4, from 143, i.e., 8.5%, to 201, i.e., 12%, of the maximum of 1664 possible wrong uses.

The total number of uses of NEC as an answer in the entire test went from 605 in the pretest to 1351 in the posttest, an absolute increase of 746 uses. This is an increase of 58 in the number of wrong uses (7.8% of the total 746 increase) versus an increase of 688 in the right ones (92.2% of the total shift). Had the 746 total increase come only from guessing due to learning just that NEC is as legitimate as yes or no, the increase in the wrong uses of NEC should have been equal to that of the right uses, i.e., 373 each, rather than the observed increase of 58 and 688 respectively. The expected posttest use of NEC based upon that assumption comes to 835 (= 462 + 373) times rightly, and 516 (= 143 + 373) times wrongly. How likely is it to obtain these observed 58 and 688 increases, if indeed all the experimental group had learned was that NEC was an acceptable alternative? How significant is the difference between the observed use of NEC in the posttest -- 1150 times rightly and 201 times wrongly, and the expected posttest use of NEC based upon the above assumption -- 835 rightly and 516 wrongly?

Had the data in table 5.17 been based upon independent observations, a binomial test would be proper for answering the first of the two questions, and a chi-square test would provide an answer for the second. Because data are based upon adding across individual students' uses of NEC, which cannot be regarded as independent events, no statistical analysis of these data with regard to those questions was found adequate. Hence, it is left to the reader to judge the significance of these results. To the experimenter it seemed exaggerated to assume that the above results were obtained even though absolutely no new understanding of the logic behind the NEC answer was gained. Therefore another hypothesis concerning the amount of guessing was tested.

Suppose only 80% of the additional posttest uses of NEC, i.e., 598 times, was due only to legitimization of this answer. Then half of the uses (299) would be wrong and the other half (299) right. This hypothesis yields an expected shift of 299 uses in the wrongly used NEC, as opposed to the observed shift of 58, which is about 5 times smaller.* The likelihood of this observation based upon the above new assumption is again to be judged intuitively. Table 5.18 shows repeated similar considerations with varying hypotheses on the percentage of the shift attributed to random use.

With the absence of appropriate statistical tests, interpretation of the results in table 5.18 were based on experimenter's intuition. It seemed reasonably safe to conclude that no more than 25% of the total NEC shift was due to learning just NEC's acceptability as an alternative answer. This may be interpreted in different ways: either each student exhibited overlearning of NEC in 25% of the items, or 25% of the students exhibited overlearning in all items (i.e., 75% of the students learned the proper use of NEC), or some other combination. At any rate, at least qualitatively there is strong evidence that a significant learning

---

*Only wrong uses are considered here. The argument for the shift in right uses is similar and gives equal results.

Table 5.18 Expected Additional Posttest Wrong Uses of NEC Based upon Various Assumptions on the Portion Due Only to Legitimization of that Answer

Observed total shift in NEC uses	Hypothesized percentage of total shift attributed only to legitimizing NEC	Expected shift in wrong uses of NEC (half of the hypothesized number	Observed shift in wrong uses of NEC	Ratio between expected and observed use of NEC
746	100%	373	58	6.4
746	80%	299	58	5.2
746	50%	187	58	3.2
746	30%	112	58	1.9
746	25%	93	58	1.6
746	20%	74	58	1.3
746	15.6%	58	58	1

occurred beyond the effect of decreasing inhibition towards NEC.*

5.5.4 Analysis of right and wrong yes/no answers with respect to hypothesized model for error prediction. As was shown in Table 5.17, there was a pretest-posttest increase by a factor of 2.2 in the total number of NEC answers on the part of the experimental group. Consequently the total number of yes/no answers decreased. Since most of the shift in NEC occurrences - 78% - was in the right-use direction, naturally the decreasing number of yes/no answers will show mainly in their wrong uses.

---

*To combine the results here with those reported in the previous section, one should keep in mind that the 58 shift in the wrong uses of NEC is the total for all students for all items. By averaging over 16 items (in which NEC could be wrong) we get 3.6 per item which is 3.5% of 104 students. The latter is the figure given in section 5.5.2.

An examination of item profiles given in table 5.3 (page 127) provides evidence that, in the pretest, students who answered yes or no to AC and DA items, i.e., answered incorrectly, in most cases did _not_ randomly guess between yes and no. In most of these items there was a pattern of preferred error which conformed with the model for error prediction stated in section 1.7.3.* Table 5.19 gives pre- and posttest item profiles for this group.**

Looking at _posttest_ undecidable items (AC and DA) in table 5.19, it is clear that in all of these items the right-answer frequency is consistently higher than the wrong-answer frequency, contrary to the situation in the pretest. However, the pattern of wrong answers did not change and, except for item 6, the pattern nicely fits the predicting model of section 1.7.3.

On the _pretest_ the majority of the students answered rightly the decidable (MP and MT) items and wrongly the undecidable ones (AC and DA). On the posttest undecidable items as well as decidable ones were answered rightly by the majority of the students. In other words they learned to separate out the undecidable case. One might question whether indeed students learned through the teaching period to distinguish between decidable and undecidable items, or whether they might have possessed this ability from

---

*The pattern is indicated in table 5.3 by the underlined numbers.

**In the posttest the control group did not change much. Therefore only the experimental group's posttest wrong answers pattern is of interest.

Table 5.19 Item Profiles for Experimental Group
Pretest and Posttest (n=104)

(Right answers are circled. Errors expected from model in section 1.7.3 are underlined.)*

Item No.	Logical Form	Negation Mode	Pretest				Posttest			
			Yes	No	NEC	Skipped	Yes	No	NEC	Skipped
1	MP	++	92	4	7	1	94	7	3	0
23		++	83	13	8	0	89	3	12	0
25		+-	31	60	13	0	5	85	13	1
30		+-	3	95	5	1	11	83	10	0
11		-+	95	1	8	0	90	5	9	0
13		-+	83	9	12	0	84	2	18	0
19		--	2	97	4	1	6	82	14	2
3		--	8	88	8	0	3	87	14	0
22	MT	++	1	97	6	0	2	84	16	2
31		++	2	89	10	3	7	88	9	0
8		+-	17	69	16	2	6	82	15	1
16		+-	3	95	6	0	14	76	14	0
5		-+	53	33	16	2	83	5	15	1
9		-+	89	1	14	0	79	10	14	1
20		--	87	13	4	0	76	17	10	1
28		--	72	26	6	0	73	16	15	0
2	AC	++	79	1	23	1	36	3	64	1
14		++	80	5	16	3	35	0	68	1
4		+-	66	1	33	4	27	2	73	2
7		+-	79	2	23	0	18	2	84	0
21		-+	42	31	31	0	4	20	79	1
27		-+	1	77	25	1	10	31	62	1
12		--	31	47	26	0	6	28	70	0
18		--	3	75	25	1	11	28	65	0
17	DA	++	5	71	27	1	5	25	74	0
10		++	6	67	30	1	0	35	69	0
6		+-	49	18	37	0	14	18	69	3
32		+-	30	37	34	3	20	9	73	2
29		-+	4	66	30	4	7	22	74	1
15		-+	8	52	44	0	2	33	69	0
24		--	54	9	41	0	21	6	77	0
26		--	82	5	17	0	17	6	80	1

*This table in fact reproduced data given previously in tables 5.2 and 5.4. The numbers here, however, are numbers of students, not percentages.

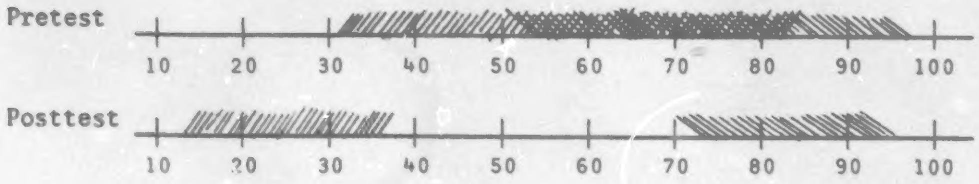
the beginning, but were unable to exhibit it due to their reluctance to choose the answer "not enough clues," for psychological reasons irrelevant to logical reasoning. If the latter is right and, as a result, undecidable items were answered at random, each of the three alternative answers -- yes, no, not enough clues -- would have had the same frequency on each AC and DA item of the pretest. Or, at the least, the yes/no answers which were wrong on the undecidable items would have had equal distribution per item. As table 5.19 shows, this was not the case. The existence of a clear wrong-answer pattern in the AC and DA subtests indicates that in tackling undecidable items students in general were not guessing but thinking. Sometimes their thinking yielded the right answer -- NEC -- other times it yielded an explainable yes/no mistake. The ratio between the frequency of right answers and that of expected errors is in favor of the expected error on the pretest and is in favor of the right answer on the posttest for all AC and DA items. The clear evidence of non-random wrong answers on the pretest, along with the pretest to posttest shift from expected error to correct answers indicate that indeed progress in the experimental-group students' conditional reasoning ability accounts for the marked gain in score.

As mentioned already, there was a noticeable difference between the number of right and the number of wrong answers for each of the pretest decidable items in favor of the right ones, but this was not so for undecidable pretest item profiles. Posttest profiles show preference to right answers in all items. These observations should be interpreted as follows: on the posttest, experimental-group students were highly capable of determining whether or not a conclusion followed necessarily from the premises, but on the pretest they failed to do so. Most often they regarded a conclusion as a necessary one while in fact it was not. This becomes clearer in comparing the patterns of right "yes" and "no" answers on MP and MT (circled in table 5.19) with patterns of expected wrong yes/no answers in AC and DA respectively (underlined). Such a comparison clearly shows that most of the wrong answers to AC and DA items can be explained by considering AC as MP and DA as MT, taking for granted that the converse of the given conditional premise holds too. Table 5.20 gives the ranges of yes/no right answers and expected errors. The ranges as given in table 5.20 overlap largely in the pretest but are distinct and far removed from each other on the posttest as can better be seen in diagram 5.8. This reinforces the conclusion drawn above about the students' change in ability to distinguish between a necessary and an unnecessary conclusion.

Table 5.20 Range of Yes/No Right Answers on MP and MT (numbers circled in table 5.19) and Expected Wrong Answers on AC and DA (numbers underlined in table 5.19)

		Right: Yes/No		Expected error: Yes/No	
		MP	MT	AC	DA
Pretest	Minimum	60	53	31	30
	Maximum	97	97	80	82
Posttest	Minimum	82	73	18	14
	Maximum	94	88	36	35

Diagram 5.8 Ranges of Yes/No Right Answers on the Decidable Part of the Test (▧) and of Yes/No Expected Wrong Answers on the Undecidable Part of the Test (▨) as Predicted by the Model in Section 1.7.3



The pretest situation supports Henkin's conjecture (section 1.7.3) about right MP and MT answers due to a fortunate mistake by "language balance" which is responsible for wrong AC and DA answers as well. It supprots also O'Brien's interpretation of mastery of MP and MT, but lack of ability on AC and DA items due to conversion or inversion errors. The posttest situation (bottom line, diagram 5.8) shows a marked difference between the ranges of yes/no right answers on decidable items (MP,MT) and the ranges of yes/no expected wrong answers on undecidable items (AC,DA). This indicates that in the posttest most decidable items were not answered rightly on the basis of some sense of language which has nothing to do with logic, but rather on the basis of valid reasoning. For if this were not the case, wrong answers due to "language balance" would occur as frequently as right ones and their ranges would not be disjoint. So even though no significant change in performance on MP and MT was observed, it is not unreasonable to assume a change in students' understanding of these forms as well as of AC and DA.

The last comment about wrong answer patterns has to do with decidable items -- MP and MT logical types. Wrong answers on those items were not as frequent as those on the undecidable items. On the average about 17% of the students gave wrong answers to an MP item and about 26% to an MT item. Comparing the pretest wrong answers on MP and MT items to the posttest ones we get the following table 5.21:

Table 5.21 Average Number of Wrong Answers Per Item in MP and MT.

	Pretest		Posttest	
	Opposite	NEC	Opposite	NEC
MP	8.9	8.1	5.2	11.6
MT	16.3	11.0	9.6	13.5
Decidable subtest	12.6	9.6	7.4	12.6

The data in table 5.21 shows a clear difference between the pretest and the posttest patterns of wrong answers on decidable items. Whereas on the pretest more wrong answers were opposite answers (yes where no is right, and vice versa) than NEC answers, the situation on the posttest was reversed. Expressing undecidability between yes and no instead of a straight contradictory answer may at first appear as a change in the right direction. However, this is not a real gain since both are wrong. It is more reasonable to attribute this change to the recognition of NEC as a valid answer, and possibly to some overlearning it as discussed in the previous sections.

5.6 Standard School Achievements and Sex Differences with Respect to the Results of the Present Study.

5.6.1 Non-correlation with S.A.T. and with sex. The Stanford Achievement Test (S.A.T.) was administered in all eight participating classes about four months prior to the pretest. The distribution of National Stanine as determined by that test, and the mean national grade-equivalent for each class were given in table 4.2. The total battery mean scores for experimental and control groups were given in table 5.1. Table 5.22 shows correlation coefficients between (i) sex, the three mathematics subtests of S.A.T., reading and science subtests of S.A.T., and (ii) pretest, posttest, and gain scores on the four logical-form subtests of the present study for the experimental and the control group. No number in table 5.22 is above .45. In fact most of them are close to zero. In particular none of the five S.A.T. subtests considered were found to be good predictors for success in the pretest, in the posttest, or for learning the experimental material as measured by the pretest-posttest gain scores. Intuitive logic seems in this case to be uncorrelated with other school topics such as mathematics in any of its standard school aspects. These results indicate that no linear relation is a good approximation for the relations between any of the five ordinary school topics studied and application of the relevant basic logical rules on an intuitive level. Such findings reconfirm the need for change in the unbalanced school mathematics curriculum discussed in section 1.1.2, or for an independent place in the curriculum for logic.

There was also no correlation between students' sex and either

Table 5.22 Correlation Coefficients Between (i) Sex, S.A.T. Subtests, and (ii) Pretest, Posttest and Gain Scores of the Present Study.

[In each square the upper left number is for the experimental group (n=104) the lower right number is for the control group (n=106)]

Test		Sex	Stanford Achievement Test				
			Mathematics			Reading	Science
			Concept	Application	Computation		
PRE-TEST	MP	.07 .07	.08 .19	.24 .22	.27 .17	.21 .10	.15 .05
	MT	.16 .08	.23 .21	.13 .21	.15 .18	.19 .26	.18 .24
	AC	-.02 .00	.39 .12	.29 .14	.35 .23	.25 .24	.24 .23
	DA	-.03 .10	.43 .27	.32 .17	.34 .40	.35 .31	.23 .30
POST-TEST	MP	.27 .21	.23 .05	.20 .09	.25 .05	.32 .00	.21 -.04
	MT	.00 .04	.29 .21	.32 .25	.30 .14	.24 .24	.24 .22
	AC	.17 -.05	.38 .31	.26 .24	.25 .42	.32 .29	.20 .39
	DA	.10 .03	.44 .39	.41 .31	.37 .43	.45 .43	.33 .43
GAIN SCORE	MP	.18 .14	.14 -.12	.00 -.10	.03 -.10	.12 -.08	.07 -.08
	MT	-.14 -.05	.03 -.03	.14 .00	.11 -.06	.03 -.06	.03 -.05
	AC	.18 -.07	-.06 .26	-.07 .14	-.14 .26	.02 .10	.07 .23
	DA	.13 -.07	-.06 .19	.03 .19	-.04 .09	.03 .18	.05 .20

pretest, posttest, or learning in the present study as measured by gain scores.

5.6.2 High, average, and low S.A.T. achievers' performance in the present study. Table 5.23 summarizes pretest, posttest, and gain-score means for experimental-group subgroups by high, average, and low national stanine levels in regular school mathematics, reading, and science. The small number of students in each of the low stanine categories makes questionable the reliability of any analysis based upon the obtained low stanine's means. However, a comparison of the high stanine subgroup with the average stanine subgroup for corresponding S.A.T. subtests shows that for all logical forms, pretest and posttest scores of the high stanine subgroup were higher than those of the average stanine subgroup. But both levels of stanine showed very similar patterns of gain scores. In fact, in AC, DA, and totals, the gains of the average stanine subgroup were higher than those of the high stanine subgroup on three out of the five S.A.T. subtests -- math-computation, reading, and science. This implies that on the posttest the average stanine subgroups for these three topics came closer to the high stanine subgroups than they had on the pretest. In all cases both high and average level students benefitted significantly from the experimental unit. The mean scores for the entire experimental group, which were analyzed in the previous sections and were found statistically significant in AC and DA, were therefore about equally influenced by the progress of above-average students and average ones. The few students whose national stanine was below average also benefited. These results encourage the teaching of intuitive logic in regular classrooms, rather than as a special topic for better students.

Table 5.23 Pretest, Posttest, and Gain Score Means for Subgroups of the Experimental Group (n=104) by Levels of Achievement on S.A.T.

Present Study \ S.A.T.		High Stanine (7,8,9)					Average Stanine (4,5,6)				
		Math Concept n=52	Math Application n=55	Math Computation n=30	Reading n=51	Science n=51	Math Concept n=48	Math Application n=45	Math Computation n=69	Reading n=49	Science n=47
MP (max=8)	Pre	6.8	6.8	7.0	6.9	6.8	6.5	6.6	6.5	6.5	6.5
	Post	6.9	6.9	7.1	7.0	7.0	6.5	6.4	6.6	6.5	6.3
	Gain	.2	.1	.1	.1	.2	.0	-.2	.0	.0	-.2
MT (max=8)	Pre	6.5	6.3	6.7	6.6	6.5	6.0	6.1	6.1	5.9	6.0
	Post	6.5	6.5	6.6	6.5	6.5	5.9	5.7	6.1	5.9	5.8
	Gain	.0	.1	-.1	-.2	.0	-.1	-.4	.0	.1	-.2
AC (max=8)	Pre	2.8	2.6	3.6	2.4	2.5	1.2	1.2	1.1	1.5	1.3
	Post	6.2	5.9	6.3	5.9	5.6	4.6	5.0	5.0	5.1	5.3
	Gain	3.5	3.3	2.7	3.5	3.2	3.4	3.8	3.9	3.6	4.0
DA (max=8)	Pre	3.4	3.2	4.0	3.2	3.2	1.8	1.7	1.8	1.9	1.6
	Post	6.5	6.4	6.7	6.4	6.2	4.8	4.9	5.2	5.0	5.2
	Gain	3.1	3.2	2.7	3.2	3.0	3.0	3.2	3.4	3.1	3.6
Total (max=32)	Pre	19.4	18.9	21.3	19.1	19.0	15.6	15.6	15.6	15.7	15.4
	Post	26.1	25.6	26.4	25.7	25.4	21.8	22.0	22.9	22.5	22.6
	Gain	6.7	6.7	5.4	6.6	6.4	6.3	6.4	7.3	6.8	7.2

		Low Stanine (1,2,3)				
		Math Concept n=4	Math Application n=4	Math Computation n=5	Reading n=4	Science n=6
Total (max=32)	Pre	12	16.2	18	15.8	18.5
	Post	21	20.3	20.8	18.5	21.33
	Gain	9	4.1	2.8	2.8	2.83

5.7 Attitudes

5.7.1 Students' attitudes.

At the end of the pretest in both the pilot study and the main study, students were given a two-item questionnaire to obtain some indication of their attitude (see appendix 7.4a). The first item asked the students to circle the names of the activities they liked and to cross out those that they did not like. Table 5.24 shows the number of students who did not cross out the various activities. Students' answers showed that the Electric Cards activity was the most popular. The Pictorial activity was the least favored for the main-study students, yet it was the second most popular with the pilot-study students. This could have been caused by the change made in this activity between the two studies: A word-puzzle page was inserted after each pictorial page. Although reactions to the pictorial pages were positive, as teachers reported, students in the main study

Table 5.24 Number of Experimental Group Students Who Indicated They Liked Each Activity (of Maximum 104 Students).

Name of Activity	Number of Students Indicating They Liked It (maximum 104)
Electric Cards	94
Pictorial Activity	17
Dominoes	69
Numbers and Their Properties	18
Playing Cards	76
Colored Light Switchboard	87
Prepare a Quiz	75

complained heavily about the amount of writing the puzzle pages required. These complaints, coming in the beginning of the teaching period, had a strongly deleterious influence on teachers' attitudes as will be discussed in the next section.

Other differences between pilot- and main-study students' attitudes towards activities were apparent with the dominoes. The pilot-study students favored this set of activities, but the main-study students reacted as if the activities required only trivial thinking. There are two possible explanations for this difference. Pilot study students were urban children of socio-economic class lower than that of main-study students, who were suburban. Pilot study students' general ability was also probably lower because it is known to be correlated with socio-economic class. The second possible reason was the difference in sequencing. The dominoes activity was presented later in the main-study teaching period than in the pilot-study teaching period. The experience main-study students gained by previous activities may have caused the activities with dominoes to appear less challenging.

The opposite happened with Numbers and Their Properties. Modifications made in this activity between the two studies in an effort to increase its power of motivation seemed to be successful. Main-study students were challenged and enjoyed it more than the pilot-study students. The activities with the colored-light switch box were attractive to many students in both studies.

The experimenter's impression was that providing such a large variety of experiences in the undecidable case gave some students the feeling that it's "more of the same" and therefore their attention wandered sometimes.

The second item on the student-attitude questionnaire was phrased: "Right after Easter vacation" (Christmas vacation, for the pilot study) "we may offer a new logic project." (Logic project was the common name for the experimental unit.) "Only those students who choose to participate in it will take it. It's up to you. Would you like to go on learning logic after Easter (Christmas)?______ What's your reason for your answer? (space)." Table 5.25 gives the distribution of yes answers, no answers, and undecided answers for the second item for the pilot- and the main-study students.

Table 5.25 Attitude Measure of Experimental Group Students by Volunteering to Go on Learning Logic After the Posttest

Would you volunteer?	Pilot study (n=49)	Main study (n=104)
Undecided	13.9%	12.6%
No	22.4%	29.4%
Yes	63.7%	58%

Here are some examples of the answers as summarized in table 5.25.

Undecided: "It matters what we do. I don't like some of the things we do, some I like. If we do the things I don't like, I won't and vice versa."

"I don't know. It was fun but sometimes it was boring."

"I did not like writing the answers, but logic was interesting."

"I am not sure. Depends if I have time."

"Maybe. I liked learning how to make the electric cards, etc., etc., But I hated the pictures." (I.e., pictorial activity which involved giving written arguments.)

No: "I didn't like to answer why."

"Because I did not like much of it. I never could tell if I was right."

"I hate it. It's hard and confusing."

"No. I just don't want it any more. I had enough of it."

"It was boring. So boring and hard."

"It would be like extra work. Thanks anyway."

"Because I'm not really learning anything, although I liked it pretty much."

Yes: "Because I liked it all."

"It's just fun."

"Most of the stuff was fun."

"Because you get to build things and get more smarter."

"It is fun and I learned new things. It takes a lot of thinking, too."

"I liked to make some of the things and it is more fun than schoolwork. P.S. It's after Easter! P.P.S When do we start?"

"I think it would be interesting because it gets you to think hard. And it's easier than what we do in school."

In general, as table 5.25 shows, the majority of the students in both studies were ready to volunteer to continue learning logic. Because their judgments were most likely based only upon their experience with logic through the present study, it was concluded that most of the students liked the experimental unit.

5.7.2 Teachers' reactions. During the pre-training sessions all experimental group teachers in the main study were very cooperative and enthusiastic. They expressed worries about students' ability to grasp the material in a short period of time* and tried to think of ways to prevent student-frustration. When the teachers were told they would be asked to fill out evaluation sheets on each activity, one said: "Oh, I'll give you miles of feedback because I know how important it is to develop such units." The investigator's impression before the beginning of the teaching period was that the teachers felt confident and capable of handling the teaching. Teachers took the test a day before their students were pretested and reported it took time but they got all but one or two items right. In their answers to items 16 and 17 on the overall-evaluation-of-the-experimental-unit questionnaire (appendix 7.4b) they filled out after the posttest was over, all teachers confirmed that their preparation was adequate. One wrote: "The sessions were sometimes too long. On the whole we learned very much and it was really enjoyable." None of the teachers thought there was a way to improve the teachers' manual so that it would be the sole source of instruction for the teacher (item 15). Each in his/her own way expressed the need for training in logic. "I enjoyed the experience of learning logic - very needed," said the most critical teacher. However, they felt they were not given enough opportunity to assist in the development or discuss the course with each other during the pretraining. "Too much 'top down'," one said.

---

*However, later the project seemed to be too lengthy to them.

The experimenter met with the teachers once a week during the teaching period. In the first of these meetings, after the first week of teaching during which "Electric Cards" and then "Pictorial Activity" were presented, the teachers seemed very disappointed by their classes' reactions. A resistance on the students' part to give written reasoning on the puzzle pages of the pictorial activities developed, and the teachers felt a drop in the level of motivation. In his answer to item 19 of the overall evaluation questionnaire, one of the teachers said: "They liked to mess with the electric cards a lot at first, but as soon as they got into the workbooks (the pictorial activity),the morale dropped considerably after the first 20 pages." Another teacher said: "They were excited at the beginning, bored in the middle, and felt competent and involved in the project at end." The experimenter's impression was that the same could be said for the teachers' change of attitude throughout the teaching period.

Two of the four teachers positively stated that they would like to teach this unit next year (item 11). "I feel that the underlying ideas of this project should be introduced throughout the elementary education. I definitely want to teach these ideas next year." "I like the innovative themes of this unit." The other two teachers expressed doubts, saying: "Maybe. I could leave out some activities and I would present it in a more integrated way. It was too isolated and lengthy." The other said: "If I do, it would be shortened and integrated into a larger program of analytic thinking skill."

The teachers felt that five weeks were too condensed or too long a period. "There were too many activities with too little time." "It's novelty wore off quickly." "Too much sameness too often."

One of the four teachers saw no possible carry-over effect. Three teachers expressed hoped for carry-over effects of the experimental unit on students' work in other parts of the curriculum (item 22 and 23). "Hopefully - they will be able to reason better in discussions and in writing." "Using these ideas I may be able to instil the abstract idea of math better." "Hopefully, they will be more aware of assumptions." One of the two teachers who answered doubtfully about teaching the unit again said here: "It was a worthy experiment. I and the kids learned a great deal."

All four teachers stated (item 14) they read the teachers' manual and needed about half an hour of preparation for many sessions. The teachers' manual itself seemed to be clear enough but too detailed, leaving too little freedom to the teacher. "I wish my expertise as a teacher could have been called on."

In answering item 21 two teachers said they were surprised at the small extent to which the composition of ability groups in regard to the experimental unit conformed to general ability. "Many kids are motivated to figure out puzzles yet are not motivated in general classroom curriculum."

Activities preferred by teachers (item 1) were: "Prepare a quiz - it was challenging and manipulative." "Pictorial activity - nonverbal, challenging like a game." "Playing cards - it demonstrated the impossible case, p and not q, best of all in my opinion." Least favored activities were: "Numbers and Their Properties - the most abstract of all." "Dominoes - I feel I wasn't prepared enough for it."

When teachers were informed of the results of their teaching they were pleased with the results. They obviously felt satisfaction. One said: "Their reasoning was so surprising." However, one of them said: "We had too little freedom. Now I feel creative again."

To summarize: teachers were appreciative of the need for a unit like the experimental one, they felt generally well prepared, and were challenged to take over the teaching of the unit. On the other hand teachers felt it was too condensed, too long, or too repetitive, even though very innovative.

CHAPTER 6

SUMMARY AND CONCLUSIONS

Chapter Overview

The first section in this chapter is a summary of the goals, procedures, and results of the present study.

The second section relates this study to previous studies, describes the weak points of the present study, and consequently suggests future investigations.

6.1 Summary of Findings

6.1.1 Summary of objectives. There were three main objectives to the present study. The investigator sought:

1. To develop a unit in conditional reasoning for the intermediate elementary grades aimed at familiarizing students with the distinction between valid and nonvalid inferences from simple conditional premises, through concrete factual or hypothetical examples, and without using an algorithmic approach, namely without a direct presentation of formal rules of inference;
2. To have elementary school teachers implement the unit in their ordinary classes as a regular part of their curriculum; and
3. To examine the effectiveness of that implementation in improving the students' performance in conditional reasoning.

The primary considerations that led to choosing conditional reasoning as the topic of the unit were: the central role that the logic of conditional sentences plays in mathematics and the rich existing psychological research in this area, which indicates the substantial need for improvement in young children's (as well as adults') conditional reasoning ability.

The study stemmed from a desire to redress the distorted view of mathematics in the elementary curriculum. This distortion is created by the current inbalanced emphasis on computational rules and on some application, but very little logical analysis and abstraction. This situation exists despite the wide recognition that school mathematics should be a major contributor to the development of the general ability to reason logically.

6.1.2 Summary of development and design. The experimental unit was developed in four cycles of teaching-revision-reteaching, with a changing role on the part of the investigator: (i) investigator's work with individual students, (ii) investigator's successive work with three small groups of students, (iii) teachers' implementation pilot study in which the investigator was present in each and every class period, and (iv) teachers' implementation main study, where the investigator paid only occasional short visits to each class.

To examine the effect of the experimental teaching, a test in conditional reasoning consisting of 32 three-choice items was developed through several field trials and revisions. Each test item was formulated with a reasonable hypothetical content designed to make sense to fourth and fifth graders in the selected population. In its final form, the test consists of 4 eight-item sets, each in one of the four logical forms, interspersed among each other:

MP (modus ponendo ponens)
MT (modus tollendo tollens)
AC (affirming the consequent)
DA (denying the antecedent)

The first two forms (MP and MT) constitute the decidable part of test, where the correct answers are either Yes or No. The other two forms (AC and DA) constitute the undecidable part of the test, where the correct answer is "Not-enough-clues."

In each eight-item, logical-form subtest there were four pairs of items, in each of which the conditional premise is in one of the following negation modes.

"+,+": No negation occurred in either the antecedent or the consequent;
"+,-": negation occurred in the consequent only;
"-,+": negation occurred in the antecedent only;
"-,-": negation occurred in both the antecedent and the consequent.

All items were presented in written form. The investigator administered the test as a group test.

The main study involved 104 students in four experimental classes in one school, and 106 students in four control classes in another school. All eight classes belonged to the same school district in the San Francisco Bay Area, California. The population in this district is predominantly of upper-middle socio-economic class. All four experimental group teachers took a twelve-hour pretraining workshop given by the investigator, and attended weekly meetings during the instruction period. Teaching took place for 30-40 minutes a session, 4-5 times a week for 23-25 sessions.

6.1.3 Summary of results.

A. The test-retest reliability of the conditional-reasoning test was .79.

B. Experimental and control group pretest performance levels were not significantly different (not even at .05 level). Pretest performance of both groups on the decidable subtests, MP and MT, was highly successful. The experimental and the control groups received on the average 83.3% and 81.1% right answers on MP items respectively, and 78.3% and 78.8% on MT. Both groups performed poorly on the undecidable subtests of the pretest -- AC and DA. Experimental and control group respective mean percentages of right answers were 24.3% and 22.1% on AC, and 31.3% and 33.3% on DA. The total pretest mean score was 54.3% and 53.8%, respectively. The groups did not differ significantly ($\alpha = .05$) in standard school achievements either. The mean score on the Stanford Achievement Test was 69.5 for the experimental group and 68.9 for the control group.

C. There was a significant difference ($\alpha = .01$) between the experimental group's and the control group's overall performance on the posttest -- 55.4%, the mean percent of correct answers for the control group (compared to 53.8% on the pretest), and 74.7%, the mean percent of correct answers for the experimental group (compared to 54.3% on the pretest).

D. Whereas on the pretest 37.5% of the experimental group students and 39.7% of the control group students had fewer than 50% correct answers, on the posttest 36.8% of the control group students performed at this level, but only 2.9% of the experimental group students remained at this level. On the other hand, only 8.7% of the experimental group students (9.4% of the control group students) had as many as 75% to 100% correct answers on the pretest. On the posttest, however, 52.9% of the experimental group had this percentage of correct answers (and only 10.4% of the control group).

E. As mentioned above, both groups were initially highly successful on MP and MT. Posttest scores on these logical subtests did not differ significantly from pretest scores. However, there was a marked change in experimental group performance on AC and DA subtests. Recall that the maximum number of correct answers on each logical-forms subtest was 8. The experimental group mean score on AC went from 1.9 (23.75%) on the pretest to 5.4 (67.5%) on the posttest, and from 2.5 (31.25%) to 5.6 (70%) on DA. The control-group pretest performance on these subtests, as well as on the decidable subtests, matched the level of the experimental group. Unlike the experimental group, however, the control group did not show significant change in the posttest.

F. Thirty of the 104 experimental students moved from the lowest pretest level to the highest posttest level on the undecidable part of the test. An additional 20 experimental-group students moved to the highest level from the medium performance level, and 32 others moved from the lowest pretest level to the medium posttest level. Altogether 82 students progressed, 20 stayed at their initial performance levels, and two regressed. About 75% of the experimental students retained their pretest level on the decidable part of the test. This also happened in the control group. Seventy to eighty percent of the control group students remained at their pretest levels on each of the four subtests, decidable as well as undecidable ones.

G. The order of difficulty of the various negation modes within each logical form differed from one logical form to the next. Within each negation mode, decidable items were consistently answered more successfully than undecidable items.

H. The expected wrong Yes/No answers on AC and on DA subtests in the pretest were found to be almost as frequent as the Yes/No correct answers on corresponding negation modes of the MP and MT subtests respectively. This conforms with the hypothesized model of error prediction (section 1.7.3). Experimental group posttest results show a vast reduction in error rate and a consistent pattern of correct answers on all four subtests respectively. Ranges of Yes/No correct answers on the decidable part of the test and of Yes/No expected errors on the undecidable part largely overlapped on the pretest -- 53 to 97 and 30 to 82 respectively. On the posttest these ranges were distinct -- 73 to 94 and 14 to 36 respectively. This indicates that between the two tests students learned to distinguish undecidable items from decidable ones.

I. Attempts to separate guess effect from learning the true meaning of the answer "Not-enough-clues" yielded the following results: 77.9% of the students in the experimental group exhibited learning of the true meaning of NEC by choosing this answer correctly on the posttest in more items than on the pretest in addition to one of the following: (1) either choosing NEC as an answer wrongly on the posttest less often than on the pretest, or (2) choosing NEC wrongly on the posttest more often than on the pretest, but the difference between the additional correct uses and the additional wrong uses was at least three.* Of the experimental group students, for 58.7% this difference was at least six. It is fairly safe to infer that at least 55% of the students learned the logical meaning of NEC. An analysis of the shift between pretest and posttest in the overall number of correct and incorrect uses of "not-enough-clues" as an answer by the experimental group yielded a total shift of 743 uses: a pretest to posttest shift of 58 incorrect use, and a shift of 688 in the correct use of NEC. It seems unreasonable to assume that such a large improvement in correct usage is attributable to a mechanism other than learning of the logical meaning of the "not-enough-clues" answer.

The percentage of experimental students choosing NEC answers per item went from an average of 7.8 on the pretest to 9.9 on the posttest for MP, and from an average of 9.4 to 12.9 on MT. In both cases, NEC is a wrong answer. Therefore an increase of 3.5 was interpreted as overlearning of the NEC answer. (Control group results show fluctuation of ±1.9% on on MP and MT which should be subtracted from the above 3.5) Now, on AC and DA, where NEC is correct, the average percentage of experimental students choosing it as an answer showed an increase of 43.6 and 39.0 respectively leaving an item-average increase of 43.1 and 35.5 which is not due to over-learning of NEC but eventually to learning its logical meaning.

---

*Not included among these 77.9% are students (19.2%) whose pretest to posttest fluctuations in both wrong and right use of NEC was in the range of ±3. They were omitted because this pattern of fluctuation was typical of the control group and therefore was attributed to random factors.

J. More than half (58%) of the students expressed a positive attitude toward the experimental unit by stating they would like to take a second similar course if offered on a voluntary basis. An additional 12.6% could not decide, and 29.4% would not volunteer. No comparison between volunteering for a further experimental unit and for any ordinary school activity was taken.

Teachers were excited in the beginning, frustrated in the middle, and felt competent and involved in the project at the end. They did feel that the teaching period was too condensed (even though also too long) and too repetitive. The teachers were, however, appreciative of the innovation and of the educational gains that they and their students made.

K. Surprisingly, no correlation was found between the learning of logic and achievement levels in standard school subjects such as mathematics (computation, application, or concepts), reading, and science. Gain scores in the experimental-group tests of high and average achievers in these traditional subjects were not significantly different. However, high achievers started out at a higher level than average achievers and thus finished at a higher level.

6.2 Relation to Previous Studies and Suggestions for Further Investigation*

6.2.1 Test scores and their various interpretations.

Ennis and Paulus (1965), McAloon (1969), Carroll (1970), and Weeks (1970), who previously investigated the teaching of conditional reasoning to various ages and ability levels, found parallel results to those of the present study. All of these studies found that the core of observed teacher-influenced change in students' performance is in their improved ability to guard against fallacious reasoning by recognizing when conclusions do not necessarily follow from given premises.

The concentration of teaching effects in this area results from the high initial performance of students on the two valid patterns of inference from conditional premises: MP and MT. As previously found by Hill (1961) and O'Brien and Shapiro (1968, 1970, 1971, 1973), there was little room for observable progress on these parts. It is therefore not surprising that teaching had a negligible observed effect on performance in these forms. However, as previously established by the above studies and others cited in Chapter 1, students initially were unable to recognize an insufficiency of data to validly infer a given conclusion. This recognition is where teaching had the most apparent effect -- in previous studies as well as in the present one.

---

*In this section indented paragraphs indicate suggestions for further investigations. The rest consists of comparative comments.

Nevertheless, the observations concerning the initial ability levels are given an alternative explanation in this study. Previous researchers attributed the definite answers of students on undecidable items to inversion or conversion errors. Namely, it was hypothesized that students interpret $p \rightarrow q$ as $p \leftrightarrow q$; in other wrods they erroneously believe that $p \rightarrow q$ implies $q \rightarrow p$, or possibly not $p \rightarrow$ not q.* This explanation presupposes the ability of the students to choose correct answers on MP and MT by logical analysis. An alternative suggestion is that a phenomenon closer to Sells' "atmosphere effect" is responsible for the observed success on MP and MT. "Language balance," as conjectured by Henkin (1974)** provides an explanation independent of logical analysis for students arriving at correct answers in both MP and MT, and also wrong answers in both AC and DA by a mechanism that has nothing to do with logical reasoning. To a certain extent, the regression observed in MP and MT in the pilot study of the present work, and the decline in MT reported by Ennis and Paulus (1965) support Henkin's conjecture. If students were consistent on MP and MT on the pretest due to sound logical analysis, how does one explain their loss of some of the ability when introduced to the undecidable cases? If it is easy to confuse them, do they really possess the initial ability to recognize necessary conclusions?

---

*The findings of the present study, as well as previous ones, show that DA is answered more successfully than AC. This may indicate, if indeed MP and MT are known in advance, that students interpret $p \rightarrow q$ as implying the inverse "not $p \rightarrow$ not q" and not as implying the converse "$q \rightarrow p$." When the inverse is considered as given, DA becomes MP and AC becomes MT. Thus the order of difficulty is consistent with the order of difficulty of MP and MT. Further investigation is needed to find out more about the source of these errors.

**See section 1.7.3.

a. Henkin's conjecture was not proved in this study, nor was it refuted. Obviously further investigation of this conjecture is needed. One possible approach was discussed in 1.7.3.

b. Because "language balance" occurs independently of the conditional connective, a study of other connectives may shed light on this matter.

6.2.2 Negation in the conditional premise. Negation in the conditional premise of an item does influence the degree of difficulty of the premise as measured by the number of students who answered a given item correctly. Findings of this study show, parallel to previous ones (Roberge 1969 and O'Brien 1972) that various logical forms are more difficult in some negation modes than in others.

In the present study there were only two items of each negation mode within any 8-item logical form subtest. Therefore, reliability of the results of negation modes within a particular logical form is compelled to be low. Consequently only the relative rank order of difficulty of the four modes is discussed here. The correspondence of these orders with O'Brien's results (1972) obtained from high school subjects and based upon three items of identical mode within each logical form is noteworthy. Table 6.1 gives this comparison. The two orders agree only in MT. However neither of the two studies should be considered as giving sound generalizable data on the problem of order of difficulty; each study is based on a very small number of items per case. Both studies, however, point out that in some cases additional negation is not necessarily a factor that increases the difficulty of an item. In fact, in DA and AC forms, items with no negation in the

Table 6.1 Order of Number of Correct Responses for Negation Modes within Logical Forms in O'Brien's (1972) Study and in Pretest for the Experimental Group (n=104) of the Present Study

MP	O'Brien's study	++ > -- > +- > -+
	Present study	-- > ++ > -+ > +-
MT	O'Brien's study	++ > +- > -- > -+
	Present study	++ > +- > -- > -+
AC	O'Brien's study	+- > -- > ++ > -+
	Present study	+- = -+ > -- > ++
DA	O'Brien's study	+- > -- > -+ > ++
	Present study	-+ > +- > ++ > --

first premise were found in both studies to be less successful than others.

The exact order of difficulty of the several negation modes within logical forms, however, calls for further investigation. The data available on this matter in previous studies and in the present one is not sufficiently reliable.

6.2.3 Implementation problems. Results reported in Weeks' study (1970) puzzled the experimenter. Second and third graders were trained solely by means of attribute blocks. The training covered the same length of time as that of the present study. According to Weeks' report, the students improved significantly on a 36-item test representing a much broader scope of logical inferences than the scope of this study. It is not clear what attitude Weeks' students developed through this period, but the present experimenters' experience with a program of more variety suggests that they must have become quite bored playing with just one manipulative. Osherson (1974) admits a developed boredom on the part of his subjects in the process of playing with two manipulatives in a logic test. The novelty of any manipulative in the present study wore off within three to four sessions. Also, no measure of learning transfer to other logical forms was taken in the present study. There was an apparent common feeling among the teachers that the teaching was too lengthy. Weeks' report of the success of his own teaching of a very broad scope of logical inferences to much younger students using only a single manipulative aid contrasts with the effort needed to be invested in the present study to achieve significant progress in a more limited scope teacher implemented unit. Weeks worked with small groups whereas this study took place in regular size classes. All these factors taken together indicate a need for further investigations in the following directions.

a. Development for fourth and fifth grades of teacher-implementable units in which a variety of logical patterns, richer than those addressed in the present study, are introduced at an intuitive level. Patterns of quantificational (first order) logic which are much more adequate for mathematical arguments, deserve wider attention at the school level.

b. Development of teacher-implementable units in intuitive logic for grade levels lower than fourth and fifth grades, possibly Weeks' students age -- second and third grade.

c. Search for efficient methods of inservice training and trials with a wider sample of teachers. The need for change in college preparation given to prospective teachers was indicated in Chapter 1. However, mathematics education cannot afford to wait for this change. An inservice-training program intertwined with the development of units for students is necessary to accomplish a widespread implementation of units in logical reasoning in the elementary schools.

d. Investigation of the transferability of training in certain parts of mathematical logic to other untrained parts.

e. Investigation into more effective ways of obtaining and maintaining young students' motivation and attention while learning logic on an intuitive basis. For example, a modification in the approach of the present study should be attempted. A gradual progress could be made through a hierarchy of prerequisites, instead of simply repeating similar experiences in a variety of situations. The sense of making progress may contribute to the student's feeling of "that's worthwhile doing."

f. Comparison of concentrated training versus training spread over time.

g. A study of the progress effected by each activity (with a single manipulative aid), versus the effect of the whole series of activities presented in this study. The purpose would be to isolate the most effective activities, and possibly to compare them with the effect of attribute block training on conditional reasoning.

6.2.4 Levels of regular school achievements and socio-economic class. C. Carroll (1970) worked with ninth-grade low achievers in mathematics in an effort to induce a change in their conditional-reasoning ability. In most of the comparisons she made, she found no significant differences between the experimental group and two different control groups. By contrast, in the present study the fifth graders who were low achievers in mathematics concepts and applications (4% of experimental group) did make significant progress.

a. This limited data should by no means be considered as more than an indication of the need for further study of the effect of introducing young, low-achieving students to logical analysis in general, and to conditional reasoning in particular.

b. The present study was implemented partially in a small middle class city, and partially in a suburban upper-middle class community. The extent to which teacher-implemented units in basic logic on an intuitive level can be effective for lower socio-economic classes should also be investigated.

6.2.5 Content effect. The desire to introduce young children to logical analysis requires a search for appropriate methods. One major consideration should be given to the content within which logic is treated. Further investigations relating to content are suggested in two different directions,

a. Despite repeated revisions, type-mate items of the present study test were not uniformly consistent. Content effect probably rests behind this inconsistency. Some unexplained findings were indicated in the analysis of test results. Enlightenment concerning content effect should occupy future research, if a non-abstract approach to the teaching of logic is to be sought.

b. The present unit includes only one activity which is directly related to mathematics. In Chapter 1 the viewpoint was expressed that logic should be an integral part of the mathematics curricula. While the one mathematically related activity of the present study may provide an example of the interweaving of logic with fifth grade mathematics, a further effort should be made to integrate the units in logic into the mathematics curricula. Isolation brings about the difficulty of transfer into mathematics, where the use of logic is necessary. Isolation also creates uncertainty as to the purpose of learning logic. Natural integration into well established school topics like mathematics may prevent these troubles, as well as balance the undue emphasis on computation in today's school mathematics. This integration might also produce a higher correlation between performance in ordinary school subjects and in logical analysis, which was found to be null in the present study.

6.2.6 More doubts about the present study. In this section additional weak points of the present study are mentioned and future research possibilities are offered.

a. The sample of students in the present study was not randomly selected. The selection of classes was determined by the teachers who volunteered and the school districts willing to cooperate. This creates doubt about the generalizability of the results in many ways. The number of teachers was too small to control for the teacher variables. Therefore generalizability of teaching effects is doubtful. Also the population of the students was a specific section of the population so generalizability to other sections is questionable.

b. Boredom and motivation as factors in student success on the tests were not studied. Observed results might give an obscured picture if students did not put enough thinking effort into answering test items despite its team competition nature.

c. The investigator in the present study took an approach of mass education. No effort was made to find out what was going on in individual minds by interviewing or by any other method. Statistical analysis of pretest and posttest results is considered by the experimenter as a very powerful approach to studies of learning and of teaching effects, but not necessarily the best way to study learning or thinking processes.

d. Is the effort worthwhile? Even if students are indeed able to learn to distinguish valid from fallacious inferences, should not such topics be postponed to a later age when students could "breeze through" it? Is there such an age? These questions are very important to answer before any attempt is made to integrate units in "intuitive logic" into elementary school mathematics curricula.

e. Neither transfer nor retention were studied this time. A comparison of transfer and retention of units such as the present experimental one and regular school curriculum might add to decision making regarding the value and the usefulness of units in "intuitive logic."

Finally, there is more unknown than known in the area of logic education. Logic education should certainly occupy mathematics educators extensively, since the ability to carry out valid logical analysis and guard against fallacious arguments is essential for the learning and understanding of mathematics.

There is much to do in this area and a long way to go. No one work can provide a complete cure for the complex problematics of mathematics education and logical analysis. In educational research, more than in other disciplines, what one may reasonably expect is slow and steady improvement. The present study is intended to be one small step in this direction.

CHAPTER 7

APPENDICES

APPENDIX 7.1

FINAL VERSION OF THE EXPERIMENTAL UNIT - TEACHERS' MANUAL AND ANSWERED STUDENTS' WORKSHEET

CONDITIONAL REASONING

An Intuitive Approach To The Logic Of Implication

Teacher's edition

(Experimental)

by

Nitsa Hadar

Lawrence Hall of Science

University of California, Berkeley

January 1975

"PERMISSION TO REPRODUCE THIS COPYRIGHTED MATERIAL HAS BEEN GRANTED BY
The Regents of the University of California
TO ERIC AND ORGANIZATIONS OPERATING UNDER AGREEMENTS WITH THE NATIONAL INSTITUTE OF EDUCATION. FURTHER REPRODUCTION OUTSIDE THE ERIC SYSTEM REQUIRES PERMISSION OF THE COPYRIGHT OWNER."

Preparation of this manual is being supported by the American Association of University Women, and by NSF grant no. PES 74/18950.

Copyright © Regents of the University of California, 1975.

Table of Contents

Title . . . . . Page

The logic of conditional reasoning 1
A word about the objectives 9
General comments to the teacher 10

Electric cards 12

Activities with dominoes 18

Pictorial activity 26
    Answers sheets 33

Numbers and their properties 65
    Answers sheets 71

Playing cards 77
    Answers sheets 83

The colored light switch box 85
    Answers sheets 91

Prepare a quiz 99
    Answers sheets 107

No part of this manual can be reproduced without permission of the Regents of the University of California, through the Lawrence Hall of Science.

List of symbols and their meanings

$\wedge$	Will stand for "and".
$\vee$	Will stand for "or" (inclusive or).
-( )	Will stand for "It is not the case that" (particularly for negating a conditional sentence).
p,q	Will stand for simple sentences (i.e., sentences which don't include the words "and", "or").
p → q	Will stand for "if p then q (a (conditional sentence).
$\overline{p}$	Will stand for "not-p" (similarly $\overline{q}$ is "not-q").
MP	Is the rule by which one infers q from p → q and p.
MT	Is the rule by which one infers not-p from p → q and not-q.
AC	Is the fallacious rule by which one invalidly infers p (or not-p) from p → q and q.
DA	Is the fallacious rule by which one invalidly infers not-q (or q) from p → q and not-p.

The Logic of Conditional Reasoning

1. What is a conditional sentence?

Given any two declarative sentences we can obtain a conditional sentence by putting the word "if" at the beginning of one sentence, the word "then" at the beginning of the other one, and joining the resulting clauses. The conditional sentence obtained in this way will read:

(sentence 1) (sentence 2)
"If......................, then....................."

For example:

sentence 1: It is raining.
sentence 2: The grass is wet.
Conditional sentence: If it is raining, then the grass is wet.

(When the two sentences are chosen randomly, the conditional sentence may not make sense.)

The first sentence forming the conditional sentence, the one that follows word "if", is called the antecedent. The second one,following "then", is called the consequent. We'll often use the letter "p" to stand for the antecedent, and "q" for the consequent. A conditional sentence will then be denoted by $p \rightarrow q$ (read: if p, then q).

2. What does a conditional sentence imply?

(a) Let's use the above example to demonstrate the general case which will be discussed later. We had: p = It is raining

q = The grass is wet

and $p \to q$ = If it is raining, then the grass is wet. Whenever this sentence is true we cannot have both of the following true: *it is raining*, and *the grass is NOT wet*. Writing $\bar{q}$ for not-q, (i.e., for "the grass is *not* wet"), the notation $-(p \text{ and } \bar{q})$ means that it is *not* the case that it is raining and the grass is not wet. Thus, the sentence $p \to q$ always implies the sentence: $-(p \text{ and } \bar{q})$.

(b) The sentence "*If it's raining*, *then the grass is wet*" excludes *only* the event: *It's raining and the grass is not wet*. No other events are excluded. In other words, the sentence: "*If it's raining*, *then the grass is wet*" leaves three possibilities open: (1) It's raining and the grass is wet; (2) It's not raining and the grass is wet; (3) It's not raining and the grass is not wet. Symbolically, if $p \to q$ is true then at least one of the following three is true: (1) p and q; (2) $\bar{p}$ and q ($\bar{p}$ means not-p); (3) $\bar{p}$ and $\bar{q}$. Using the word "or" in its inclusive meaning, we could write: $p \to q$ implies $(p \text{ and } q)$ or $(\bar{p} \text{ and } q)$ or $(\bar{p} \text{ and } \bar{q})$.

(c) Any conditional sentence $p \to q$ not only *implies* each of the sentences $-(p \text{ and } \bar{q})$; $(p \text{ and } q)$ or $(\bar{p} \text{ and } q)$ or $(\bar{p} \text{ and } \bar{q})$, but is actually *equivalent* to each of these sentences. This means that given any of these three sentences, we may confidently conclude the other two. Still differently, it means that all three of them say exactly the same thing.

(d) Let's consider our example again. We said: If it's raining, then the grass is wet. If this is so, can it ever happen that the grass is not wet and it's raining? No. We've seen in (a) that this is impossible. In other words: if the grass is not wet, then for sure it's not raining. The last sentence is again a conditional sentence. It is different from our original one, but it expresses exactly the same thing. We call the new sentence $\bar{q} \rightarrow \bar{p}$, the contrapositive of the original sentence $p \rightarrow q$.

3. What's not implied by a conditional sentence?

(a) Once again let us suppose: If it's raining, then the grass is wet. We saw in 2(b) that this sentence does not imply anything about the wetness of the grass when it's not raining. In particular, the conditional sentence: If it is not raining, then the grass is not wet, does not follow from the original sentence. The grass may be wet even though it is not raining (someone may have watered it, for example).

In general, the sentence $p \rightarrow q$ does not imply $\bar{p} \rightarrow \bar{q}$. (Notice the difference between $\bar{p} \rightarrow \bar{q}$ and the contraposition $\bar{q} \rightarrow \bar{p}$ which is implied by $p \rightarrow q$.)

(b) Similarly, $p \rightarrow q$ does not imply $q \rightarrow p$. Back to our example, the sentence: "If it's raining, then the grass is wet," does not imply that "If the grass is wet, then it is raining"!

(As we said before, the grass may be wet for many other reasons.)

4. Two true rules of inference

(a) Given a conditional sentence along with its antecedent, we can confidently conclude its consequence. Example:

Given (I) If it's raining, then the grass is wet

(II) It's raining

Conclusion: The grass is wet.

This inference pattern is called "Modus Ponens". We'll abbreviate it by MP. Symbolically, MP is the following rule:

$$\begin{array}{ll} (\mathrm{I}) & p \to q \\ (\mathrm{II}) & p \\ \hline & q \end{array}$$

where the line separates the assumptions from the conclusions.

(b) Given a conditional sentence: *if p, then q*, we've seen in 2(d) that its contrapositive: *if not-q, then not-p*, necessarily follows. Therefore, given any conditional sentence p → q and *knowing* that not-q, we are lead to conclude that not-p must be the case.

An example will clarify it:

Given (I) If it's raining, then the grass is wet

(II) The grass is *not* wet

Conclusion: It is not raining. For if it is raining, then, by (i) the grass must be wet. But (ii) contradicts it, and hence, it cannot be raining.

This inference pattern is called "Modus tolendo tollens", abbreviated by MT. Symbolically it allows us to infer in the following way.

$$\begin{array}{ll} \text{Given (i)} & p \rightarrow q \\ \underline{\text{(ii)}} & \underline{\text{not-}q} \\ & \text{not-}p \end{array}$$

where, again, the horizontal line separates the givens from the conclusion.

5. Two fallacies

(a) We had no difficulty to accept that if it's raining, then the grass is wet. Suppose now we look out through the window and see that it's not raining. Is the grass wet?- We can't tell. The information we possess does not suffice for any definite answer for that question. The grass may be either wet or not when it's not raining. The only thing we do know is what happens to the grass when it is raining. Nothing is said about it when it's not raining.

People sometimes are tempted to reach a definite conclusion when a conditional sentence is given along with the denyal of its antecedent. Their conclusion is fallacious. Their logical pattern that yields this fallacy is called "Denying the antecedent", abbreviated by DA. Symbolically:

Given (I) $p \rightarrow q$
(II) not p
no valid conclusion

(b) Another logical trap in conditional reasoning is the acceptance of the consequence (abbreviated AC). Formally

Given (I) $p \rightarrow q$
(II) q ,
no valid conclusion exists

But, people sometimes tend to conclude p in this case. In relation with our previous example, given that the grass is wet, people very often conclude that it is raining. This conclusion cannot validly be drawn from the two premises because there is nothing in the premises to prevent the grass from being wet while it's not raining.

6. Negation mode of a conditional sentence

As stated above, a conditional sentence $p \rightarrow q$ has two parts, the antecedent p and the consequent q, which in themselves are sentences. Negation may occur in each part. For example, p may say "His hands are dirty", and q may say "He doesn't get dinner". The conditional sentence will then read "If his hands are dirty, then he doesn't get dinner". In this case p does not include negation and q does. In general, there are four cases (negation modes):

(1) p and q don't include negation (e.g., If it's raining, then the grass is wet).

(2) p includes negation and q doesn't (e.g., If our car is not fixed yet, we'll stay home during the weekend).

(3) p does not include negation and q does (e.g., If he is busy, then he will not agree to join us).

(4) Both p and q include negations (e.g., If a child is not 5, then he does not go to public school).

Whatever was said above about conditional sentences holds for any such sentence in any negation mode. It should be noted, though, that when p includes a negation, not-p doesn't. Let's take the example in case (3) above to demonstrate the four logical inference patterns:

MP:

(i) If he is busy, then he will not agree to join us. . . $p \to q$

(ii) He is busy. . . . . $\underline{\quad p \quad}$

Therefore: He will not agree to join us. . . . . q

AC:

(i) If he is busy, then he will not agree to join us. . . $p \to q$

(ii) He does not agree to join us. . . . . $\underline{\quad q \quad}$

No conclusion can be drawn. . . . . ?

He may be tired (not busy) or he may be busy, we can't tell.

Note: In the last example (AC) sentence (ii) affirms the consequent of (i). The negation in it is not a part of the logical structure.

DA:

(I) If he is busy, then he will not agree to join us. . . . $p \rightarrow q$

(II) He is not busy. . . . . not-p

No conclusion can be drawn. . . . . ?

His being busy does not imply that he will join us. He may still agree not to join us for many other reasons.

MT:

(I) If he is busy, then he will not agree to join us. . . . p [illegible] q

(II) He agrees to join us. . . . . not-p

Therefore he is not busy. not-p

Otherwise, according to (I), he would not have agreed to join us.

Note: In the last example (MT) sentence (II) negates the consequent of sentence (I) even though it does not include any negating word. Formally we should have written for not-q(sentence (II)): He does-not-not agree to join us. Double negation in spoken language is usually replaced by its positive equivalent as appears in sentence (II) above.

To summarize, the logical structure can't be determined by sentence (II) itself. It is the relation of it to either parts of the conditional sentence (I) that determines the logical structure and the existence of a valid conclusion.

A word about the objectives

The objective of the unit on conditional reasoning is to improve students' ability to properly apply MP and MT and to avoid the two fallacies DA and AC, when the conditional sentences are of a simple kind, namely they don't include the words: and, or (i.e. p,q in p → q are not compound sentences), and the content expressed by the conditional sentence is either factual or hypothetical, but familiar and makes sense.

Research studies of children's conditional reasoning indicate that negation add an additional obstacle in properly applying MP and MT and in recognizing insufficiency of the premises to yield any conclusion. It is among the objectives of the following unit to teach students how apply MP, MT and how to avoid AC and DA where the conditional sentence takes any one of the four negation modes (see 6 above).

In the following pages you'll find many suggestions for activities. Most of them are group activities, some of which can be conducted by the students themselves while the teacher is working with another group. Teacher who prefers the frontal teaching style may use these activities for the whole class. In general, the activities are aimed at increasing the students' ability to interpret precisely what a conditional sentence says and what it excludes, to rephrase a conditional sentence in the contrapositive way, to apply MP and MT and to avoid AC and DA. In all cases, all four logical patterns are presented for every conditional sentence. At the beginning, negation is omitted from the conditional sentences but later on all four negation modes take place.

General comments to teachers

1. Questions in the students' book may be used for paper and pencil work or for oral discussion between the teacher and the class or among peers.

2. Most of the answers are given in full in the teacher's manual. Children are not expected to give them in this exact way.

3. It is recommended that the teacher will consistently use the "four-fold contingency table" as a representation of any conditional sentence in discussion.

4. The concept: conditional sentence should be used from start without any definition. Just say something like: "Let's consider this conditional sentence ...." or: "This is a conditional sentence, what does it tell us?", etc.

5. There is no need to complete all the suggestions given in any activity in one session. When students seemed to get bored (or before this even starts) they should be switched to another activity.

6. Numbers in parentheses refer to questions in students' workbook.

7. The teaching process involves manipulative aids, games, contests and pencil-paper work. They all should be carried out to the

point the students are still interested. There is no need that every student will do everything. In any activity the main importance is to lead toward improvement on the abilities mentioned in 1, above.

8. The use of shortcuts is a preparatory step for using letters as variables. It may facilitate the generalization of the four logical structures, MP, MT, AC, DA, which is the main goal of this unit. It is not the purpose of this unit to teach the algorithm for applying correctly MP and MT. Students are expected to think about the given content and its implication. However, it is the purpose of this unit to provide experience in conditional reasoning out of which syntactical generalization will emerge. Teacher is requested not to teach the algorithm. If and when a student discovers it - it will be a great achievement. Each student should discover it by himself. Teaching the algorithm will prevent this individual discovery. It may stop the reasoning process and turn it into a mechanical routine.

9. Page numbers in students' edition are denoted by s (e.g., S-1, S-2, etc.) Comments and answers are typed in italics in students' pages.

10. The following two pages are the first ones in the students' workbook. It is recommended that students read it at an early stage of learning the unit.

11-a

S-I

To the student -

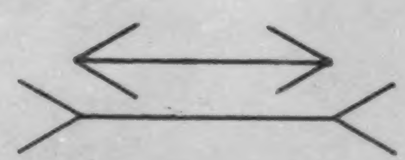
People use their eyes to see, their ears to hear, their tongues to taste, their hands to touch and their noses to smell. These are the five major senses we have. But there is another sense which is very strong in human beings, the sense of being reasonable. When we argue with someone we use this sense to provide reasons which will convince him, and also help us judge his arguments. These reasoning and judgmental processes originate in the brain through the use of logic.

Here are two puzzles. In trying to solve them, figure out which senses you use:

1. Are the two line segments equal in length? If not, which one is longer?

Line a

Line b



Answer ____

2. Tom told Frank about some rules in his family. He said: "In our house, if the TV is on, the radio must be off". One day Frank went to visit Tom in his house. When he came in, the radio was off. Was the TV on?

Answer ____

If your answer to the first puzzle was: line b is longer than line a, then your eyes misled you. If your answer to the second puzzle was: yes, then your logical sense misled you. (Measure and see that line a and line b are the same length. Think again about puzzle 2. The right answer to it is - maybe. The TV could be on, but it could also be off).

11-b

S-II

You see, our senses sometimes "trick" us. We have to train them not to. The material presented in this workbook is part of a unit aimed at helping you develop your ability to use your logical sense properly, namely to draw the right conclusions, and to be aware of possible mistakes. The work in this booklet will relate to other class activities such as games, contests, and oral discussions with your classmates.

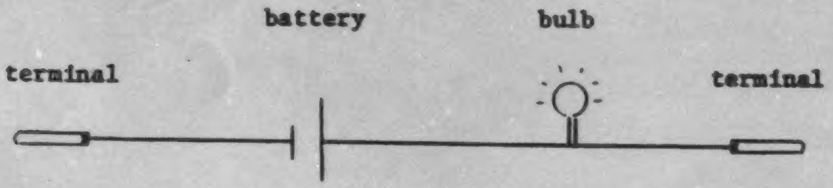
Follow your teacher's directions carefully, and hopefully you'll find this unit both enjoyable and useful.

Electric Cards

Objective: To enrich children's experience in applying MP, MT and avoiding AC, DA through a self-conducted activity with an immediate feedback.

Materials:

The Tester (Each class needs 6 testers)



The Cards

On each of the 250 cards there is a puzzle which consists of 2 clues and a question, and four metal buttons labeled as follows:

Yes, for sure	No, certainly not	Not enough clues
○	○	○
	○	

To test an answer you put one terminal of the tester on the lower button and the other terminal on any of the three labeled buttons. The cards are wired so that the bulb will light only when the right answer button is connected to the lower button.

Duration: Each of the following activities is planned for about 30-40 minutes. Each is repeatable by changing the sets of cards given to the students.

Administration:

Activity 1. Teacher-class electric card game. After presenting the tester and an electric card, the teacher plays the role of the reader in Activity 2 (see below), and works with the whole class as one team of that activity. Teacher should make sure that the students know how to write down for each card the card number and the answer next to it, on a separate line; e.g., #32. Yes

#147. No

After the first card with "not enough clues" answer was checked, show the students the shortcut, n.e.c., to stand for these words.

Activity 2: 5-6 teams of 5-6 students in each. Each student needs a pencil and paper. They write their name on top.

Each team will get twice as many cards as there are students in that team. Cards will be piled at the center of the team's table, face down. Students take turns being the reader.

The reader takes a card from the center of the table, and announces its number , which everyone writes in the first line. Then, the reader reads the question (and nobody yells the answer!). Each team member, including the reader, writes his answer next to the card number. The reader calls: "Who said yes?" "Who said no?" "Who said not enough clues?" Team members answer by raising hands.

When two different answers are found, the reader asks for reasoning and leads a discussion to settle the disagreement. He may need to reread the question before he lets each student give his argument. When a unanimous answer is reached by the group, the reader checks the right answer by the electric tester. If the group's answer was wrong, they refer to the teacher for an explanation, (or try to justify it on their own). If the group's answer was right, the student next to the reader becomes the leader for the next card.

Alternatively: captain checks the answer right after getting the answers from the students and students earn a point for right answers. The one whose total number of points is the highest, is the winner. This alternative has the disadvantage that students give their answers impulsively without being required to have an explicit reason.

Activity 3: A group of students may go to the electric cards in any stage of the course. Teacher can give them a set of cards according to a special interest:

a) Only MP.

(All 4 negation moods or just one negation mood, e.g., no negation in any part of the conditional sentence.)

b) Only MT.

c) Only DA.

d) Only AC.

e) Sets of four questions for which the first clue is identical.

(These cards have four successive numbers 1-4, 5-9, etc.)

f) Specific negation mood (all four logical forms or one, e.g., MT).

g) Cards that carry puzzles which have indefinite answers (DA or AC).

h) Cards which have definite answers (MP or MT).

Note: In each case students will be asked to compare the cards and to find differences and commonalities among them.

The cards are numbered and the following table shows the logical type and the negation mood for each question.

	MP	AC	DA	MT
$p \to q$	9,25,29,33,37, 45,49,53,61, 69,81,97,105, 125,157,229 Answer: Yes	10,26,30,34, 38,46,50,54, 62,70,82,98, 106,126,158, 230 Answer: NEC	11,27,31,35, 39,47,51,55, 63,71,83,99, 107,127,159, 231 Answer: NEC	12,28,32,36, 40,48,52,56, 64,72,84,100, 108,160,182, 232 Answer: No
$p \to \overline{q}$	1,5,21,65,73, 77,85,93,101, 117,149,169, 173,177,185, 189,193,197, 201 Answer: No.	2,6,22,66,74, 78,86,94,102, 110,118,150, 170,174,178, 186,190,194, 198,202 Answer: NEC	3,7,23,67,75, 79,87,95,105, 111,119,151, 171,175,179, 187,191,195, 199,203 Answer: NEC	4,8,24,68,76, 80,88,96.104, 112,120,152, 172,176,180, 188,192,196, 200,204 Answer: No
$\overline{p} \to q$	41,129,133, 141,145,161, 205,209,213, 217,221,233, 245,254,261, 265 Answer: Yes	42,130,134, 142,146,162, 206,210,214, 218,222,234, 246,258,262, 266 Answer: NEC	42,131,135, 143,147,163, 207,211,215, 219,223,235, 247,259,265, 267 Answer: NEC	44,132,136, 144,148,164, 208,212,216, 220,224,236, 248,260,264, 268 Answer: Yes
$\overline{p} \to \overline{q}$	13,17,57,89, 113,121,137, 153,165,181, 225,237,241 249,253 Answer: No	14,18,58,90, 114,122,138, 145,166,182, 226,238,242, 250,254 Answer: NEC	15,19,59,91, 115,123,139, 155,167,183, 277,239,243, 251,255 Answer: NEC	16,20,60,92, 116,124,140, 156,168,184. 228,240,244, 252,256 Answer: Yes

Note: 1. The letters p, q denote simple sentences without negation.

2. The right answer to any AC, DA question is NEC. The right answer is yes to MP in $p \to q$, $\overline{p} \to q$ moods and to MT in $\overline{p} \to q$ and in $\overline{p} \to \overline{q}$ moods. The right answer is no to MP in $p \to \overline{q}$ and in $\overline{p} \to \overline{q}$ moods and to MT in $p \to q$ and in $p \to \overline{q}$.

Activity 4: Electric cards can also be used for team contests of several types. In this case the tester will not be used by the team. The cards will just serve as a source of puzzles.

(a) Objective: Practice in applying MP; MT and avoiding AC; DA. Each team gets 8 cards (preferably two of each logical type). The leader of the team reads each puzzle and the team discusses it and gives an answer which is recorded by the team secretary. When they finish, the teams exchange cards so that each team answers all the questions. When teams finish they turn in the answer sheets and the cards to the teacher (or a neutral student) and he/she checks the answers. A team scores one point for any right answer (and possibly loses a point for a wrong answer). The winner is the team whose total score is the highest.

(b) Objective: The syntactical structure of MP; MT; DA; AC. Electric cards are only a source of conditional sentences. Caution should be taken in checking the answers. See example below.

Teacher reads the first clue and the answer. Students in each team discuss the possible second clue. The first team member who raises his hand gets permission to suggest what the second clue is. If his answer is wrong the other team gets a point. if his answer is right his team gets a point. Reasoning should be encouraged.

Note: When the answer announced by the teacher is: "Not enough clues," there are two possible second clues. Each one of them is a right answer! If a team can suggest both, it deserves an extra point. Example: Teacher says: "If it's a eucalyptus, then it's an

evergreen tree." What's a second clue and a question for which the answer is: "Not enough clues?"
Student 1: It is not a eucalyptus; is it an evergreen tree?
Student 2: It's an evergreen tree. Is it a eucalyptus?
Both of these students are right.

The above is a preparaton for "Prepare a Quiz" activity.

(c) Same as (b) but now teacher gives the second clue and the answer. (This time answers will vary because the first clue includes the second one as a part of it, so again electric cards are only a source of sentences and caution is needed in checking the answers.)

(d) Teacher gives each team a list of 5 conditional sentences. The team should construct as many puzzles as possible (20 is the maximum, but this obviously is not to be mentioned by the teacher). The team that succeeds in writing more puzzles is the winner. The invented puzzles can now be transferred to the other team for solution and another contest.

Activity 5: Students should be encouraged to invent their own puzzles, preferably in sets of four, with an identical conditional sentence. A new electric card can easily be made up for any good invention. (This is, in fact, "Prepare a Quiz" activity. See this activity for details.)

Activities with Dominoes

Objectives: a) (p → q) implies: not (p and not-q).

b) (p → q) implies: (p and q) or (not-p and q) or (not-p and not-q).

c) p → q is different than q → p.

Materials: 8-10 complete domino sets (2 for each team).

2 demonstration charts or a set of domino cards on a magnetic board (see description below).

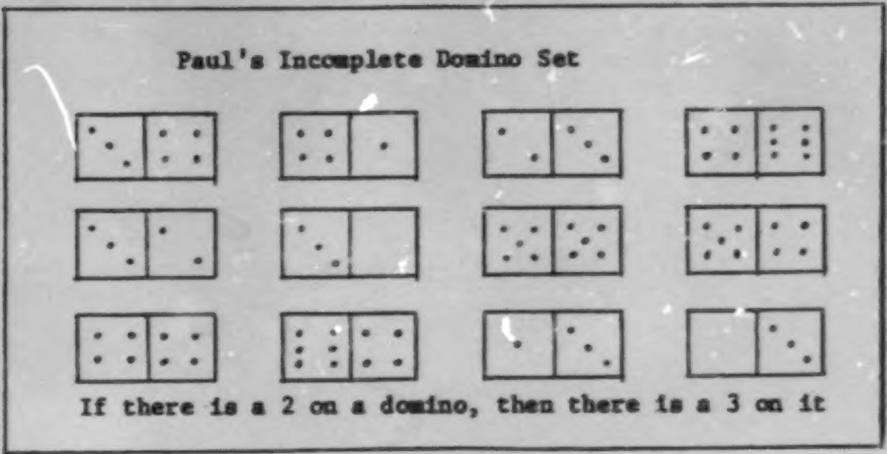
Duration: 3-4 sessions.

Administration:

Activity 1. Motivation

(a) Teacher reads or tells the following story:

Paul likes to play dominoes, but he never puts them back in their box afterwards. Therefore, he keeps losing his dominoes. One day when his friend, Peter, came over to play with him, they realized that none of Paul's domino sets was complete. Paul and Peter put all Paul's dominoes on the table face up to see which dominoes were missing. While doing so, Paul said: "Look Peter, in my dominoes, if there is a 2 on a domino, then the other number is 3." Here are the dominoes Paul had. (Post a chart): -


Paul's Incomplete Domino Set
If there is a 2 on a domino, then there is a 3 on it

Demonstration Chart #1
(See note after Part (c))

What would you say to Paul?

Is his statement correct?

(Students will discuss and will reach the conclusion that, yes, Paul was right. Arguments may be as follows: All the dominoes which show a 2 show a 3, too; there is no domino with a 2 on it which does not have a 3 on it; a domino either shows a 2 and a 3 or it does not show a 2 at all.

If difficulties arise check each domino in the chart and lead a discussion like:

T: Does this domino show a 2?

S: Yes.

T: Does it show a 3?

S: Yes.

T: Does it agree with the rule?

S: Yes.

T: Does this (the next one) domino show a 2?

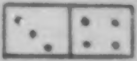
S: No.

T: Does the sentence say anything about a domino with no 2 on it?

S: No.

T: Does it disagree with our sentence?

S: No.

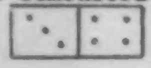
Note: A student may point at a domino like  as one which disagrees with the sentence. Teacher should persuade him through questions that this is not the case, since the sentence only tells us what happens when there is a 2 on a domino. It does not make any statement about dominoes with 3 on them, particularly it does not say that these must have a 2 on the other side.

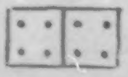
(b) T: Peter, after approving Paul's rule said: "Paul, your dominoes obey another rule." Paul didn't let Peter finish his sentence. "Wait," he said, "don't tell me, let me think about it for a second."

Can you see what Peter might have had in mind?

S: "If there is a 6 on a domino, then there is a 4 on it, is another correct rule.

Also - "If there is a blank on a domino, then there is a 3 on it."

These correct rules will be proved through checking all 12 dominoes in the chart. Any other suggestions of the form: If there is a ... on a domino, then there is a ... on it, too" will be disproved by pointing out a counter-example or, in children's language, a better term may be: a domino that denies the sentence." For example, the statement "If there is a 3 on a domino, then there is a 2 on it" is false for the above incomplete set (see demonstration chart #1) because, among others, the domino  disagrees with it.

For any suggestion of a rule try to "turn around" the sentence and ask whether the flipped over one is a true rule. For example, suppose a student suggested: "If 6, then 4," which proved to be correct. Teacher will ask: "What about-- If 4, then 6?" This is false. Proof: "  is in the set."

(c) Repeat Part (b) above with a new incomplete domino set.

For example, which rules are true for the following incomplete set?

Children may now come up with some more complicated rules like: If there is a 3 on a domino, then there is an even number on it, too. (Try its converse - if there is an even number on a domino, then there is a 3 on it, which happens to be false. Notice: both "If 5, then 0" and "If 0, then 5" are true rules.") This is in fact an introduction to Activity 2, see below.

Note about the demonstration charts: Teacher may find it useful to have a white felt board with black domino cards which can be prepared from a black cardboard with holes punched for the dots. Another idea may be to attach a magnet strip to a real dominoes and use a metal board for the demonstration.

Activity 2. Group Game: Guess Paul's Rule

4-6 players.

At least one, preferably two, complete domino sets for each team. Captain of the group needs a paper with the team members' names on it to write down the points each one earns.

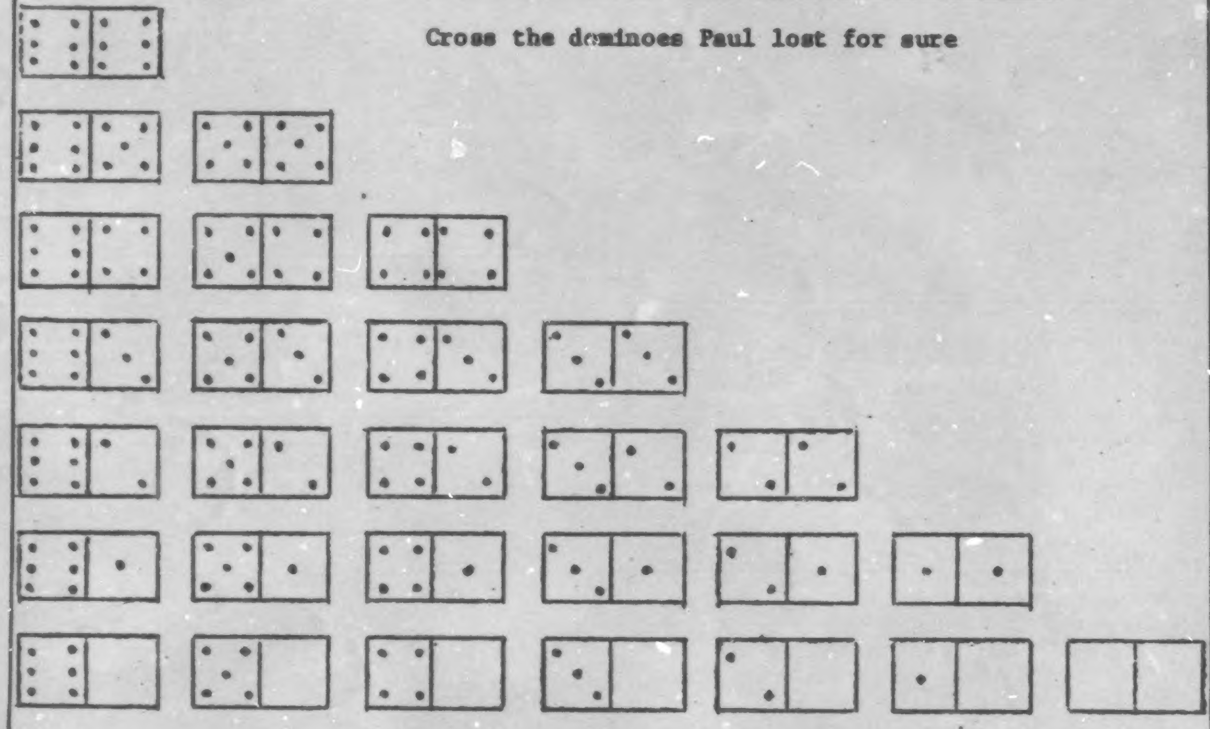
Students take turns being Paul. The student who is Paul chooses a statement of the form "if there is a ... on any of my dominoes, then there is a ... on it." He puts on the table, face up, some or all of the dominoes which obey his "rule." (An easy way for him to do it is to exclude from the complete set, all of the counter examples, as we'll do in Activity 3.) The other players have to guess "Paul's" rule. Each time a player proposes a statement, all the players check the set to see if this statement happens to be true for "Paul's" set. If it is correct, the student who suggested it earns a point. If the proposed rule is disproved, the player who pointed out a domino that denies it, earns a point. Surely, there may be more than one true rule for that set. The player who guesses the rule Paul really had in mind earns an extra point. If after 5 trials no one guessed the real rule, Paul announces his rule, and he earns an extra point. If by chance after Paul announces his rule, it is found wrong for his set, the student who discovers it first earns a point.

Activity 3. Paul's Lost Dominoes (Teacher-Class Game)

In the student's book there are a few pages with pictures of a complete dominoes set like this:

Paul's rule for his incomplete domino set is ______________________

____



Cross the dominoes Paul lost for sure

Demonstration Chart #2

Teacher will prepare a similar demonstration chart (or use felt board or magnetic board).

a) We'll start as usual by a teacher-whole-class activity. Announce the rule for Paul's incomplete set (see suggestions below). Write it down above the demonstration chart. Call students one at a time to cover off the demonstration chart with a blank card a domino which Paul lost for sure.

b) After the students got the idea, you may want to conduct a team contest: Each student will work on his own chart. Teacher announces the rule. Students write it down in the place left for this purpose above each chart. Then they cross the dominoes Paul lost for sure, according to that rule. Each student will score one point for each correctly crossed domino. Individual scores will sum up to team's score.

Here are some suggestions for rules: (examples 5, 6, are harder and may be omitted.)

1) If there is a 2 on one half, then there is a 4 on the other half. (Paul must have lost all the dominoes with 2 on one half which have 0,1,2,3,5, or 6 on the other half.)

2) If there is 4, then there is 2. (Write in short 4 → 2) In this case the following dominoes must be crossed out: all the dominoes with 4 on them and 0,1,3,4,5,6, on the other half. NOTICE the difference between the previous case and this one.

3) 1 → 5 (lost for sure: 1-0,1-1,1-2,1-3,1-4,1-6. Others, may or may not.)

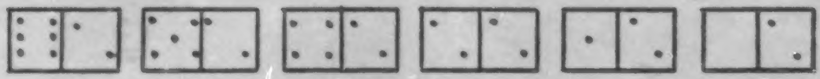
4) 5 → 1 (lost for sure: 5-0,5-2,5-3,5-4,5-5,5-6. Others may or may not.)

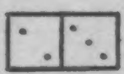
5) 5 → no-2 (lost: 5-2).

6) No-2 → 5 (lost: 0-0,0-1,0-3,0-4,0-6,1-1,1-3,1-4,1-6,
3-3,3-4,3-6,4-4,4-6,6-6)

Make the students notice the difference between their answers to 3 and to 4.

For any rule of the form "if there is an X on a domino, then there is a Y on the other half," there are 6 dominoes that are lost for sure. We know nothing about all the other dominoes. E.g., for 2 → 3 the ones that Paul for sure does not have are:

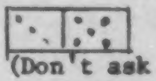


Any other domino may or may not be lost. (In particular  does not have to be in a set for 2 → 3!) As a result of that, for any conditional sentence of the above form, there are many incomplete domino sets that satisfy it.

Students should be encouraged to say a general sentence for the lost dominoes, instead of naming them one by one, e.g., in the above example, we know for sure that Paul lost all the dominoes that have a 2 on one half, but do not have a 3 on the other.

Students will need a new picture for each turn.

Note: Instead of asking a student to give his reasons for a wrong answer, lead the student to see that he is wrong by asking some more questions. For example: The rule for the dominoes Paul has 5 → no-2.

S:  (This is wrong! This one may be in Paul's set.)
T: (Don't ask why, instead ask:)
Does it have a 5 on it?
S: Yes.
T: Does it have a 2 on the other side?
S: No.
T: So, it has a 5 on it and does not have 2 on the other side. Does it meet the rule 5 → no-2?

S: Yes
T: So, did Paul lose it?
S: Not necessarily.
T: We are looking for those he lost for sure! Try another domino.

Notice: 2 → 3 does not mean that the 2 must appear on the left side of a domino that obeys that rule.

Pictorial Activity

Objectives: 1. To have children learn that $(p \rightarrow q)$ implies not (p and not-q)

2. To have children practice MP, MT, AC, DA

Duration of activity: 3-4 sessions

Materials: 18 pages with several pictures and a conditional sentence on each. Not all pictures on a given page agree with the conditional sentence on that page. Students' task is to find those pictures which disagree.

Each of the first ten pictorial pages is followed by a puzzle page on which there are four questions, one of each logical type: MP, MT, AC, DA. All four questions on a puzzle page are based upon the conditional sentence on the previous page.

The pictorial pages are arranged in an increasing order of difficulty. Pages 34-39 have conditional sentences close to children's everyday experience. Pages 40-47 have a more hypothetical nature. On pages 48-51 the conditional sentences include negation. Page 52 has no picture which disagrees with the sentence. All first ten pictorial pages (Part (a)) contain particular sentences. Pages 57-64 (Part (b)) contain universal sentences.

Pages 54, 55 are discovery pages (see answer sheets).

Conducting the Activity:

Part (a)

Each child gets a copy of these pages. Teacher will conduct the work on the pictorial pages, 34, 36, 38,...52, frontally, with the whole class (see teaching cues).

After each picture page students will work individually on the corresponding puzzle page (35, 37, 39,...,53). During this time teacher will deal with individuals, trying to lead those who wrote a wrong answer to recognize their mistakes (see teaching cues below).

Pages 54,55 are discovery pages on which children can summarize their findings. Teacher should conduct a discussion of these findings.

Part (b)

Each child will get a copy of these pages. A whole class discussion of pages 57-60 lead by the teacher, will precede individual work on pages 61-64 (see comments on answer sheets).

Teaching Cues

Pages 34,35 (S-5,S-6)

T: (Hand each student a copy of Part (a), pp. S-4 - S-26). Today we have some interesting pictures to work with. Take two minutes to browse through the pictures. See what they are about.

S: (Have a short free time to look at the pictures.)

T: Look at page S-5. (Read the instruction on top. Ask a student to read the conditional sentence at the bottom.) What do you see in these pictures? (For each picture encourage descriptions which use the words of the conditional sentence at the bottom of that page. This will be done by getting students to say for each picture: 1) what Jim is doing; 2) what time it is; 3) whether or not Jim is eating; 4) whether or not it is 7:30.)

T: I can see two pictures in which the time is 7:30, and two in which the time is not 7:30. Can you find the two pictures in which it is 7:30?

S: Pictures #1 and #2.

T: Which of these pictures disagrees with the conditional sentence?

Note: For each conditional sentence $p \rightarrow q$ it is a good habit to first separate the pictures in two parts: those in which p, the antecedent, is true(e.g., it is 7:30) and those in which p is not-true (e.g., it is not 7:30). Then, concentrate first on those where p is true (it is 7:30) and find whether or not q is true in them (Jim is eating).

Pictures in which p is true and q is not true (it is 7:30, and Jim is not eating) do not agree with $p \rightarrow q$ because $p \rightarrow q$ means that whenever p is true q will also be true.

The pictures in which p is not true (e.g., in our case, those where it is not 7:30) are irrelevant to the sentence $p \rightarrow q$ because the sentence does not discuss this case (it tells us nothing about what Jim is doing at any time other than 7:30). These pictures therefore do not disagree with the sentence, or in other words, they do not contradict anything that is said in the sentence.

S: Because it's 7:30, but Jim is not eating.

T: Very good.

Note: Each answer should be very carefully related to the sentence. There are two parts to each conditional sentence and for each picture teacher should discuss whether the first part is true or not it in, whether the second part is true or not in it and, as a result, whether the picture disagrees with the sentence. (It disagrees only when p is true and q is false.)

T: Very good. Any other picture which disagrees with the sentence?

S: #3. (Wrong answer)

T: (Ask the student to read the sentence again if you think it's necessary for refreshing his memory. Please, do not let him feel in your voice or in your face that you doubt his answer is right. We'll try to lead him to discover it.) Why do you think it disagrees with the sentence?

S: Because it is 1:30 there, and Jim is eating.

T: What does the sentence say Jim is supposed to do at 1:30?

S: It does not say anything. We don't know what he is supposed to do.

T: Does it say he is not supposed to be eating at 1:30?

S: No.

T: Does it say he is supposed to be eating at 1:30?

S: No.

T: Does picture #3 disagree with the sentence?

S: No.

T: Picture #4

S: No.

T: How come?

Note: Discuss the fact that the sentence does not tell us anything, not only about 1:30 but about any time other than 7:30. In particular, it's not impossible that Jim will eat at times other than 7:30.

It is easier to decide on the pictures that disagree with the sentence than on those which agree with it, because it is somewhat against the natural tendency to say that the picture in which it is 1:30 and Jim is playing baseball agrees with the sentence: If it is 7:30 then Jim is eating. (Since it does not disagree with this sentence, we say it agrees with it!) The terms correct and wrong are even more misleading than agree and disagree. It is therefore recommended to avoid them. We only discuss the question whether or not each picture disagrees with the conditional sentence. So always ask: "Does this picture DISAGREE with the sentence?" and not: "Does this picture agree with this sentence?" Even though this question is a negative one, it is the easier one to answer in our case.

Underlining p, the condition part of the sentence, may help the students.

S: Picture #2 does not agree with the sentence.

T: Why not?

S: Because Jim is playing. (Incomplete answer)

T: He is playing in picture #2, that's right, but this does not tell me why it disagrees with out sentence. Our sentence does not talk about playing at all. And it does not say he should not play, right? So, why does picture #2 disagree with our sentence? (If no answer is offered, ask: Is Jim eating?)

S: Because Jim is not eating. (This is a better answer, yet it is still incomplete.)

T: What's wrong with that? Does the sentence say he must be eating?

S: Yes. (Wrong! The sentence says at 7:30 Jim must be eating!)

T: Read the sentence.

S: If it is 7:30, then Jim is eating.

T: (Use your voice to stress the first part.) Oh, it says that if it is 7:30, then Jim is eating. Can you give a complete reason now, why picture 2 disagrees with this sentence?

T: Now cross out picture #2 and work quietly on the question on page 1a. (Teacher circulates around to check answers, to help children with reading, or other difficulties, and providing guidance for children who got wrong answers, by referring them back to the pictures on page 1 (see answer sheet and note after discussion of pages 36, 37).

Pages 36,37 (S-7,S-8)

T. We are going to page S-7. Mark, read the conditional sentence please. Jack, describe what you see in picture 1. (Insist on answers that use the words raining or not-raining, Carol wears her boots, Carol does not wear her boots.) Who can tell us in his own words what the sentence is all about?

S: It tells us that Carol wears her boots on rainy days.

T: Which pictures show its raining? Which of them disagree with the sentence? Why? (Similar to discussion of page 34). When there is no rain does she wear her boots, according to our sentence?

S: No. (Wrong answer.)

T: Is the sentence saying she isn't?

S: Yes. It says: If it's raining, then Carol wears her boots. So ir it's not raining, she doesn't. (Wrong argument.)

T: You said correctly that we know for sure she wears her boots whenever it rains. Read the sentence again and tell me what does the sentence say about days that are not raining?

S: Nothing.

T: So, could she wear her boots when it's not raining? Does the sentence allow her to do it or forbid it?

S: Yes. She may do it.

T: What picture shows an impossible event that can never happen. according to the conditional sentence?

S: #3.

T: Any other?

S: No.

T: O.K. Cross off picture #3. Let's work on page S-8 quietly for a few minutes.

Note: A puzzle page will be worked right after the pictorial page. Refer the students back to them in any puzzle a child has difficulties with.

E.g., puzzle #1, p. 37: If it is raining, then Carol wears her boots. It is raining. Is Carol wearing her boots? The answer is yes, for sure. On page 36 we crossed picture #3 off because it is raining there and Carol doesn't wear her boots, which disagrees with the sentence. The sentence doesn't allow that. She must wear her boots in the rain according to that sentence.

Puzzle 3, p. 37 asks about Carol wearing her boots or not on a not-rainy day. Back on page 36 both pictures #2 and #4 describe a not-rainy situation, one in which Carol wears her boots, the other in which she doesn't. Both do not disagree with the sentence. She may or may not have her boots on in not-rainy weather. There are not enough clues to decide.

Puzzle #2, p. 37: Here Carol wears her boots. Is it raining? Again, back to page 36 pictures #1 and #3 show Carol wearing her boots, but in one case it is raining and in the other it is not. Neither of these pictures were crossed off as disagreeing with the sentence. So, they both show a possible event. We don't have enough clues to decide.

Puzzle #4, p. 37: Carol doesn't wear her boots. Is it raining? Look at page 36. We crossed out the picture where it is raining, and Carol doesn't have her boots on. This is an impossible event according to the question. Carol can go without her boots on only when it is not raining. So, the answer to this puzzle is no.

Pages 38 (S-9) and on

See comments on answer sheets and follow the general scheme described above.

PICTORIAL ACTIVITY

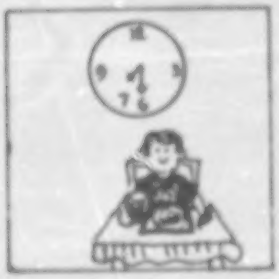
Part (a)

Teacher's edition

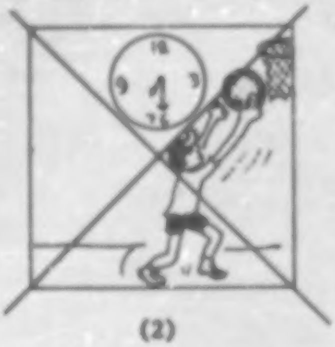
Comments and answers are typed in italics.

In each picture below Jim is doing something at a certain time. Read the conditional sentence at the bottom of this page and underline its condition part. Write the number of each picture which disagrees with it.

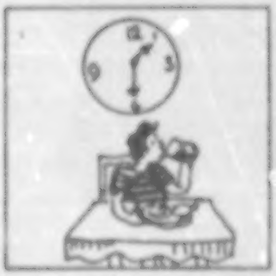
2



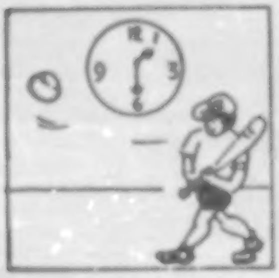
(1)



(2)



(3)



(4)

If it is 7:30, then Jim is eating.

What do you see in the pictures? (Jim is eating; Jim is not eating; it is 7:30; it is not 7:30). About what time is the sentence talking? In which pictures is it 7:30? Which of those disagree with the sentence?

Read, think, and answer:

1. a) If it is 7:30, then Jim is eating.

   b) It is 7:30.

   Is Jim eating? *Yes*

   Why? *(b) says it is 7:30, so according to (a) Jim must be eating.*

2. a) If it is 7:30, then Jim is eating.

   b) Jim is eating.

   Is it 7:30? *Not enough clues (NEC).*

   Why? *Jim may be eating at other times too. (Refer your students back to pictures 1, 3, where Jim is eating but the time is not necessarily 7:30)*

3. a) If it is 7:30, then Jim is eating.

   b) It is not 7:30.

   Is Jim eating? *NEC*

   Why? *Jim may be eating at other times too. (Refer your students back to pictures 3,4 where it is not 7:30 but Jim may or may not be eating.)*

4. a) If it is 7:30, then Jim is eating.

   b) Jim is not eating.

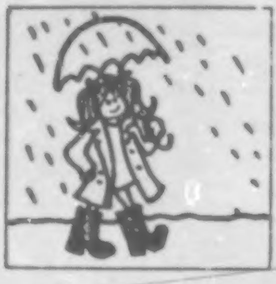
   Is it 7:30? *No.*

   Why? *(b) says Jim is not eating. If it was 7:30, then by (a) Jim would eat, but he is not. So, it is not 7:30. (That's why picture 2 was crossed off.)*

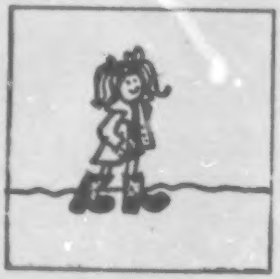
S-7

Read the conditional sentence at the bottom of this page. Some pictures on this page disagree with this sentence. Write the number of each picture which disagrees with it.

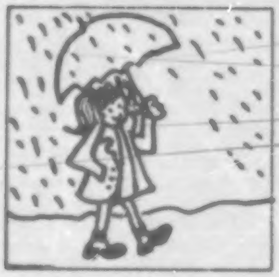
3



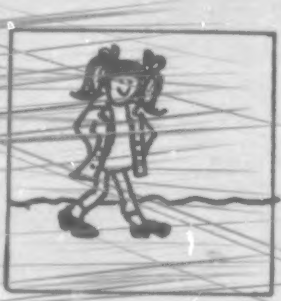
(1)



(2)



(3)



(4)

If it is raining, then Carol wears her boots.

What do you see in the pictures? (It is raining; it's not raining; Carol wears her boots; Carol does not wear her boots.) What is the subject of the conditional sentence? (It is raining. Make student underline this part.) In which pictures is it raining? Which of those disagrees with the sentence? What about pictures where it doesn't rain? (They at least do not disagree!!!)

S-8

Read, think, and answer:

1. **a) If it is raining, then Carol wears her boots.**

   b) It is raining.

   Does Carol wear her boots? *Yes*

   Why? *(b) says it's raining, in which case (a) guarantees that Carol wears her boots.*

2. **a) If it is raining, then Carol wears her boots.**

   b) Carol wears her boots.

   Is it raining? *NEC*

   Why? *Carol may wear her boots even when it does not rain. There is nothing in the sentence (a) to prevent her from doing it, (pictures 1,2) or to force her to do so.*

3. **a) If it is raining, then Carol wears her boots.**

   b) It is not raining.

   Does Carol wear her boots? *NEC*

   Why? *Even though it does not rain, Carol still may have her boots on, however she does not have to wear them (see pictures 2,4).*

4. **a) If it is raining, then Carol wears her boots.**

   b) Carol does not wear her boots.

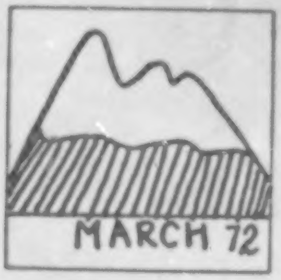
   Is it raining? *No*

   Why? *By (a) we know that if it was raining, Carol would have had her boots on. But (b) tells us she does not. So, it cannot be raining. (That's why picture 3 was crossed off.)*

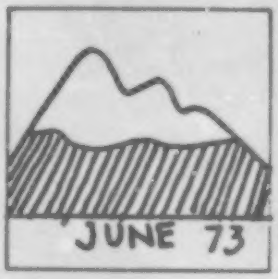
S-9

Read the conditional sentence at the bottom of this page and write the number of each picture which disagrees with it.

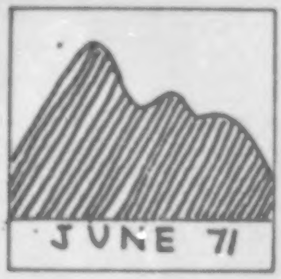
4



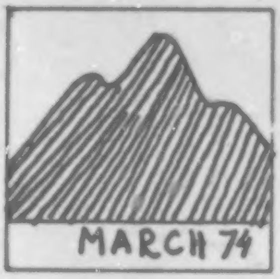
(1)



(2)



(3)



(4)

If it is March, then there is snow on this mountain.

What month of the year does this sentence discuss? Underline "it is March". What does the sentence say about that month? Find the pictures which show March. Does any of them disagree with the sentence?

What is common to pictures 2 and 3? (It is not March, in both.) What's the difference between them? Does any of them disagree with the sentence (No). Why? (See question 3.)

Read, think, and answer:

1. a) If it is March, then there is snow on this mountain.

   b) It is March.

   Is there snow on this mountain? *Yes*

   Why? *That's what (a) tells us about March.*

2. a) If it is March, then there is snow on this mountain.

   b) There is snow on this mountain.

   Is it March? *NEC*

   Why? *Even though we know for sure that in March there is snow on that mountain, there still may be snow on it any other month. (See pictures 1,2).*

3. a) If it is March, then there is snow on this mountain.

   b) It is not March.

   Is there snow on this mountain? *NEC*

   Why? *The sentence (a) does not tell us anything about months other than March. It may or may not be snow on that mountain (pictures 2,3).*

4. a) If it is March, then there is snow on this mountain.

   b) There is no snow on this mountain.

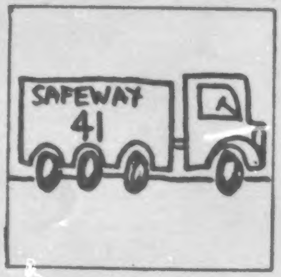
   Is it March? *No.*

   Why? *In March there is snow on that mountain (by (a)). So it cannot be March when there is no snow on it as (b) informs us.*

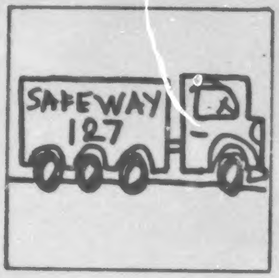
S-11

Read the conditional sentence at the bottom of this page and write the number of each picture which disagrees with it.

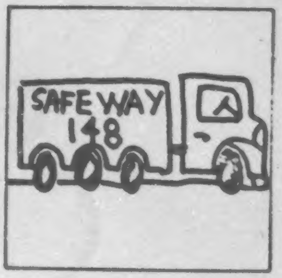
3, 4



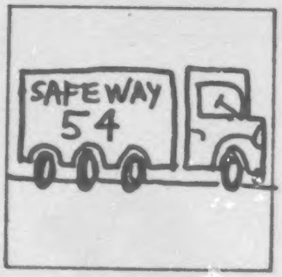
(1)


SAFEWAY
127

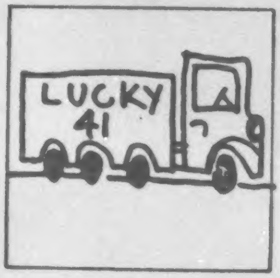
(2)


SAFEWAY
148

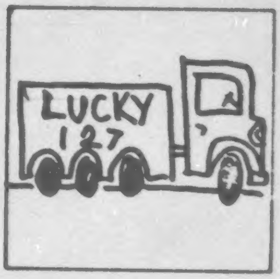
(3)



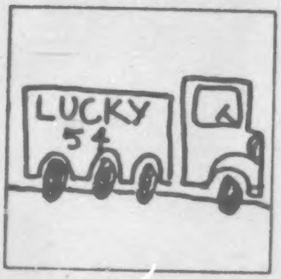
(4)



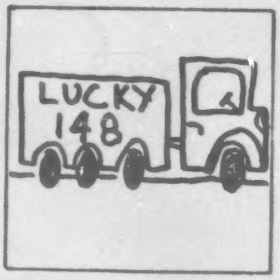
(5)



(6)



(7)



(8)

If it is a Safeway truck, then it has an odd number on it.

Explain: In one little town people discovered that the sentence below is true.

Review the notion of odd and even numbers if need occurs.

Make student recognize and underline the subject of the sentence (Safeway trucks).

Discuss the fact that we know nothing about the numbers on Lucky trucks. They may be odd or even.

S-12

Read, think, and answer:

1. a) If it is a Safeway truck, then it has an odd number on it.

   b) This is a Safeway truck.

   Does it have an odd number on it? *Yes*

   Why? *That's what (a) tells us.*

2. a) If it is a Safeway truck, then it has an odd number on it.

   b) This truck has an odd number on it.

   Is it a Safeway truck? *NEC*

   Why? *Not only Safeway trucks may have odd numbers. (Refer back to pictures 1,2,5,6)*

3. a) If it is a Safeway truck, then it has an odd number on it.

   b) This truck is not a Safeway truck.

   Does it have an odd number on it? *NEC*

   Why? *Any truck which is not a Safeway truck may carry either an odd or an even number, if it has a number at all.*

4. a) If it is a Safeway truck, then it has an odd number on it.

   b) This truck does not have an odd number on it.

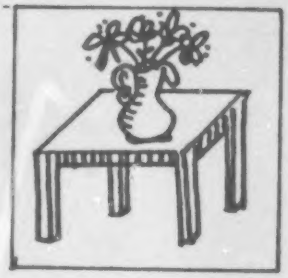
   Is it a Safeway truck? *No.*

   Why? *Safeway trucks, by (a), must have an odd number, so this truck cannot be a Safeway one since it does not have an odd number on it (by (b)).*

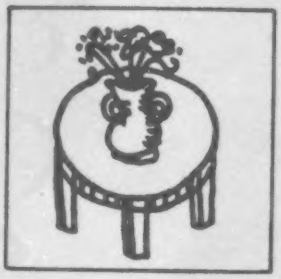
S-13

Read the conditional sentence at the bottom of this page and write the number of each picture which disagrees with it.

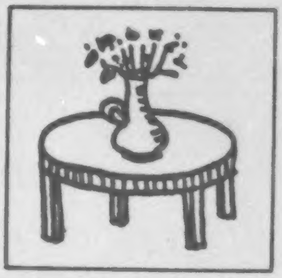
4



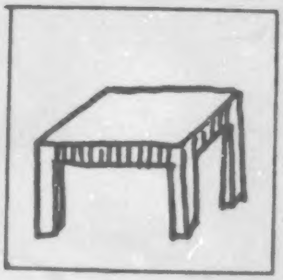
(1)



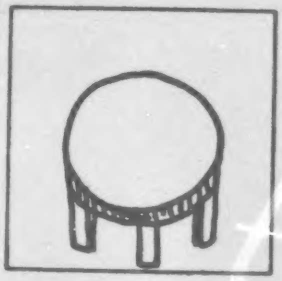
(2)



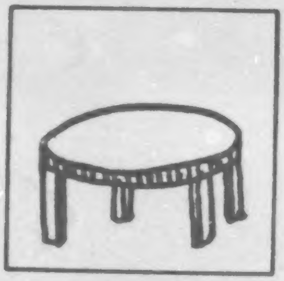
(3)



(4)



(5)



(6)

If a table has a square shape, then there are flowers on it.

Only tables with a square shape which do not have flowers on them disturb the truth of this conditional sentence. The sentence tells us nothing about any other shapes with regard to them having flowers on, or not.

Underlining the condition part of the sentence can help.

Read, think, and answer:

1. a) If a table has a square shape, then there are flowers on it.

   b) This table has a square shape.

   Are there flowers on it? *Yes*

   Why? *That's what (a) tells us for situations described in (b).*

2. a) If a table has a square shape, then there are flowers on it.

   b) There are flowers on this table.

   Does this table have a square shape? *NEC*

   Why? *Refer back to pictures 1, 2, 3 which show different shapes of tables with flowers.*

3. a) If a table has a square shape, then there are flowers on it.

   b) This table does not have a square shape.

   Are there any flowers on it? *NEC*

   Why? *Refer back to pictures 2, 3, 5, 6 which show not-square tables with and without flowers.*

4. a) If a table has a square shape, then there are flowers on it.

   b) There are no flowers on this table.

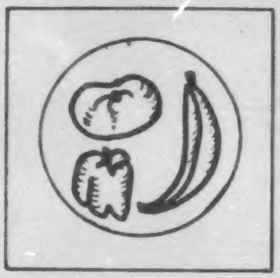
   Does this table have a square shape? *No*

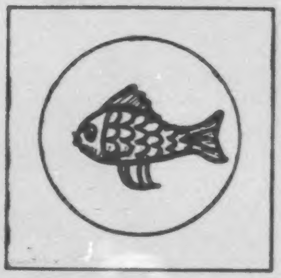
   Why? *That's why picture 4 was crossed off. A table cannot be square shaped and have no flowers on it (by (a)).*

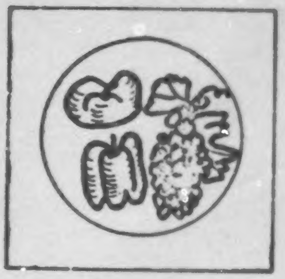
S-15

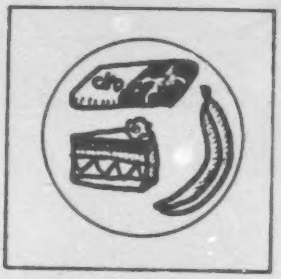
In each picture below there is a plate with some food. Read the conditional sentence at the bottom of this page and write the number of each picture which disagrees with it.

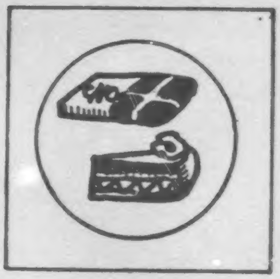
5, 6

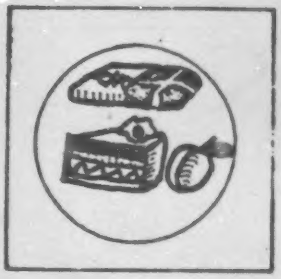

(1)


(2)


(3)


(4)


(5)


(6)

If Mom put a cake in the plate, then she put a banana in it.

Notice: A plate with a banana does not have to include a cake. It is a plate with a cake that has to include a banana!

S-16

Read, think, and answer:

1. a) If Mom put a cake in the plate, then she put a banana in it.

   b) Mom put a cake in Jim's plate.

   Did she put a banana in it? *Yes*

   Why?

2. a) If Mom put a cake in the plate, then she put a banana in it.

   b) Mom put a banana in Janet's plate.

   Is there a cake in it? *NEC*

   Why? *Refer back to pictures 1, 4.*

3. a) If Mom put a cake in the plate, then she put a banana in it.

   b) Mom did not put a cake in Jack's plate.

   Did she put a banana in it? *NEC*

   Why? *Refer back to pictures 1,2,3.*

4. a) If Mom put a cake in the plate, then she put a banana in it.

   b) Mom did not put a banana in Jill's plate.

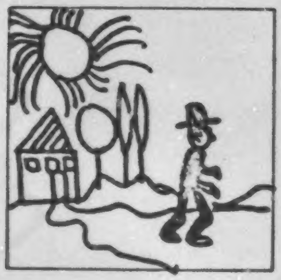
   Is there a cake in it? *No*

   Why? *There cannot be a cake in a plate without a banana in it!*

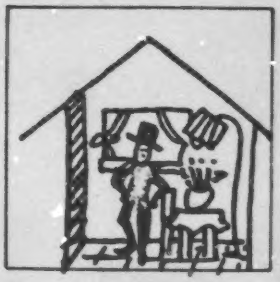
S-17

In each picture below a man named Kevin stays either in his house or outside Read the conditional sentence at the bottom of this page write the number of each picture which disagrees with that sentence.

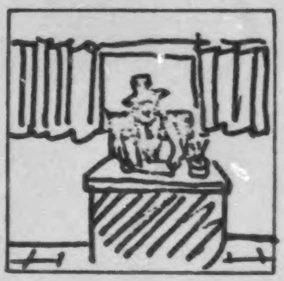
2,3



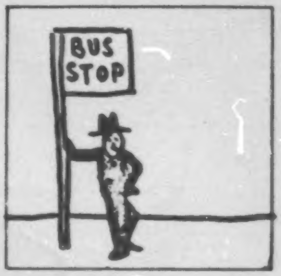
(1)



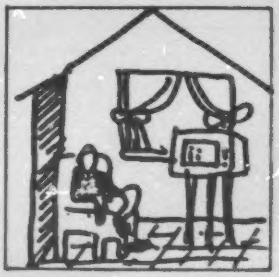
(2)



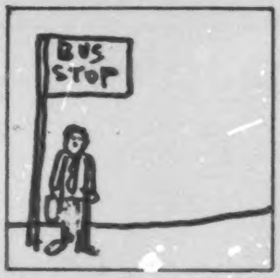
(3)



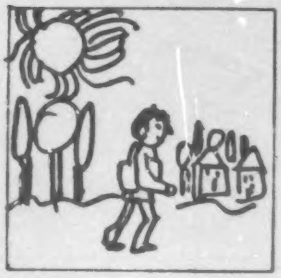
(4)



(5)



(6)



(7)

If Kevin is wearing his hat, then he is outside.

Read, think, and answer:

1. a) If Kevin is wearing his hat, then he is outside.
   b) Kevin is wearing his hat.
   Is he outside? *Yes*
   Why?

2. a) If Kevin is wearing his hat, then he is outside.
   b) Kevin is outside.
   Is he wearing his hat? *NEC*
   Why? *See pictures 1,4 and 6,7.*

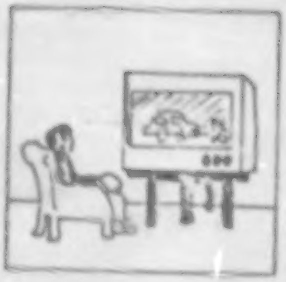
3. a) If Kevin is wearing his hat, then he is outside.
   b) Kevin is not wearing his hat.
   Is he outside? *NEC*
   Why? *See pictures 5 and 6,7.*

4. a) If Kevin is wearing his hat, then he is outside.
   b) Kevin isn't outside.
   Is he wearing his hat? *No.*
   Why?

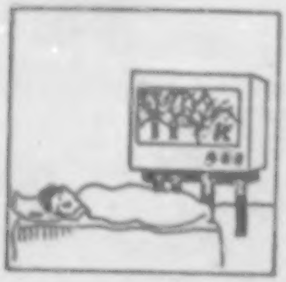
S-19

In each picture you see a man doing something, while the TV is either on or off. Read the conditional sentence at the bottom of this page and write the number of each picture which disagrees with it.

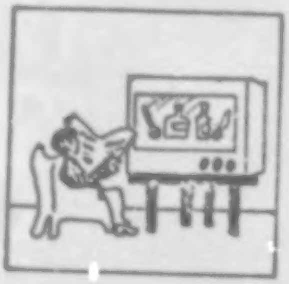
3



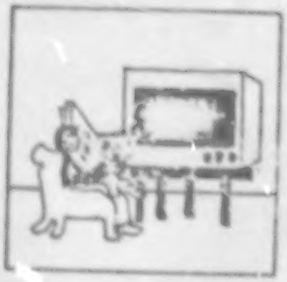
(1)



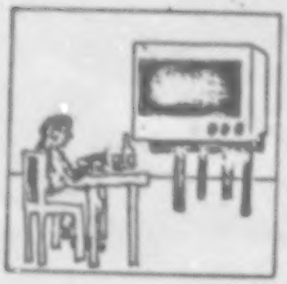
(2)



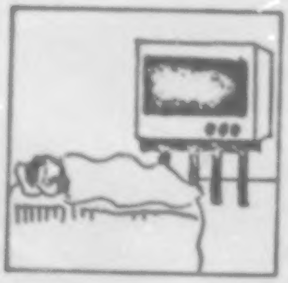
(3)



(4)



(5)



(6)

If the TV is on, then the man is not reading.

This is the first conditional sentence, with negation, in this activity. Student will as usual be called to describe the pictures in terms of: TV is or is not on; the man is or is not reading. Concentrate on those pictures where TV is on and find the picture which does not agree with the sentence which is the one where the man is reading while TV is on. Pictures 4,5,6 do not disagree with the sentence!

BEST COPY AVAILABLE

S-20

Read, think, and answer:

1. a) If the TV is on, then the man is not reading.

b) The TV is on.

Is the man reading? No

Why?

2. a) If the TV is on, then the man is not reading.

b) The man is not reading.

Is the TV on? NEC

Why? Refer to pictures 1,2 and 5,6.

3. a) If the TV is on, then the man is not reading.

b) The TV is not on.

Is the man reading? NEC

Why? See pictures 4 and 5,6.

4. a) If the TV is on, then then the man is not reading.

b) The man is reading.

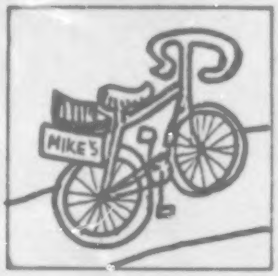
Is the TV on? No

Why?

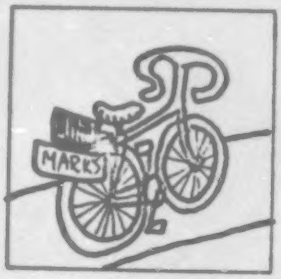
S-21

In each picture below there is a bicycle with or without a back seat. Read the conditional sentence at the bottom of this page and write the number of each picture which disagrees with it.

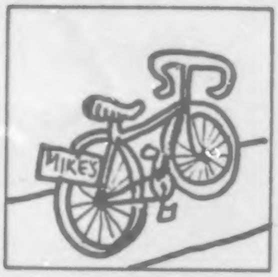
1



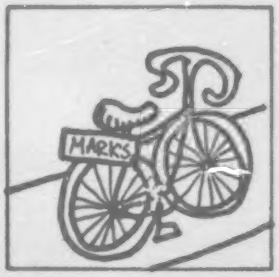
(1)



(2)



(3)



(4)

If the bicycle has a back seat, then it is not Mike's.

S-22

Read, think, and answer:

1. a) If the bicycle has a back seat, then it is not Mike's.
   b) This bicycle has a back seat.
   Is it Mike's? *No*
   Why?

2. a) If the bicycle has a back seat, then it is not Mike's.
   b) This bicycle is not Mike's.
   Does it have a back seat? *NEC*
   Why? *See pictures 2 and 4.*

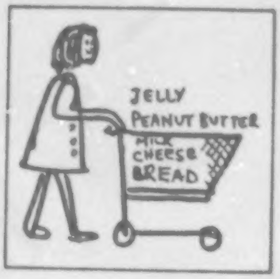
3. a) If the bicycle has a back seat, then it is not Mike's.
   b) This bicycle does not have a back seat.
   Is it Mike's? *NEC*
   Why? *See pictures 3 and 4.*

4. a) If the bicycle has a back seat, then it is not Mike's.
   b) This bicycle is Mike's.
   Does it have a back seat? *No*
   Why?

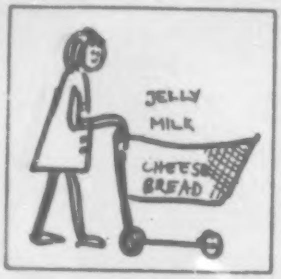
S-23

In each picture below is a complete list of things that the lady has bought. Read the conditional sentence at the bottom of this page and write the number of each picture which disagrees with it.

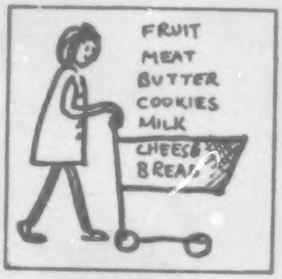
no picture disagrees.


JELLY
PEANUT BUTTER
MILK
CHEESE
BREAD

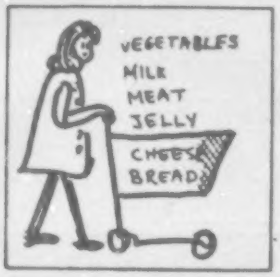
(1)


JELLY
MILK
CHEESE
BREAD

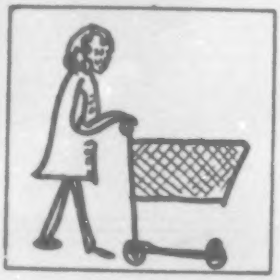
(2)


FRUIT
MEAT
BUTTER
COOKIES
MILK
CHEESE
BREAD

(3)


VEGETABLES
MILK
MEAT
JELLY
CHEESE
BREAD

(4)



(5)

If the lady buys peanut butter, then she buys jelly.

S-24

Read, think, and answer:

1. a) If the lady buys peanut butter, then she buys jelly.

   b) The lady bought peanut butter.

   Did she buy jelly? Yes

   Why? ______

2. a) If the lady buys peanut butter, then she buys jelly.

   b) The lady bought jelly.

   Did she buy peanut butter? NEC

   Why? ______

3. a) If the lady buys peanut butter, then she buys jelly.

   b) The lady did not buy peanut butter.

   Did she buy jelly? NEC

   Why? ______

4. a) If the lady buys peanut butter, then she buys jelly.

   b) The lady did not buy jelly.

   Did she buy peanut butter? No

   Why? ______

Discovery

1. Look back at each puzzle page (S-6, S-7, ... S-24). Find what is common to all four questions on each puzzle page.

Same clue (a).

The question is the same in questions 1,3 and in questions 2,4.

The answer to 2,3 is always NEC.

2. What are the differences among all four questions on each puzzle page?

Clue (b) differs

3. Can you find anything in common to question number 1 on all the puzzle pages?

Clue (b) is the condition part of clue (a). The answer is always either yes or no.

4. Can you find anything in common to question number 2 on all the puzzle pages?

Clue (b) is the second part (the result part) of clue (a). The answer is always NEC.

5. Can you find anything in common to question number 3 on all the puzzle pages?

Clue (b) is the negation of the condition part of clue (a).

The answer is always NEC.

6. Can you find anything in common to question number 4 on all the puzzle pages?

Clue (b) is the negation of the result part (the second part) of clue (a). The answer is always no. (It will be sometimes yes, in these cases, but there is no example of that kind in the previous pages. When the condition part of clue (a) includes negation, the answer to this type of question will be yes.)

PICTORIAL ACTIVITY

Part (b)

Teacher's edition

Comments and answers are typed in italics.

In each frame below results of two games are reported: California vs. Texas, and California vs. Michigan. Read the conditional sentence at the bottom and write the numbers of each frame which disagrees with that sentence.

2,3,8,9

What is the subject of this sentence? Underline it. In which frame or frames does California beat Texas? In which of those does California beat Michigan? Which ones disagree with the conditional sentence?

What does the conditional sentence tell us about games in which California does not beat Texas? (Nothing.)

California	10
Texas	3
California	15
Michigan	9

(1)

California	12
Texas	8
California	3
Michigan	15

(2)

California	9
Texas	6
California	15
Michigan	15

(3)

California	3
Texas	10
California	15
Michigan	9

(4)

California	8
Texas	12
California	3
Michigan	15

(5)

California	6
Texas	9
California	15
Michigan	15

(6)

California	10
Texas	10
California	15
Michigan	3

(7)

California	13
Texas	10
California	9
Michigan	15

(8)

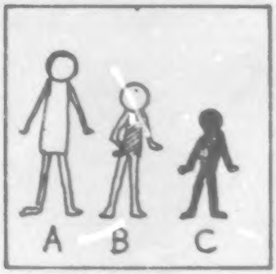
California	10
Texas	9
California	15
Michigan	15

(9)

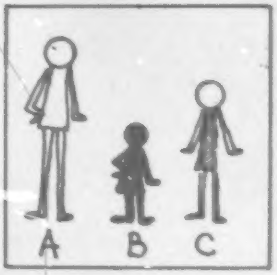
If California beats Texas, then California beats Michigan.

In each picture below there are three people. Read the conditional sentence at the bottom of this page and write the number of each picture which disagrees with it.

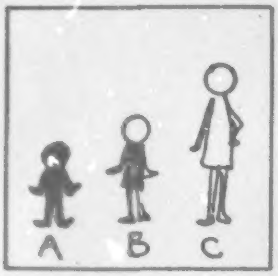
2, 7



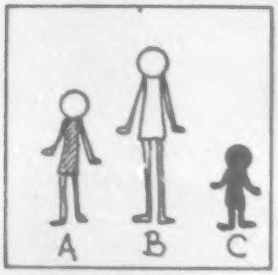
(1)



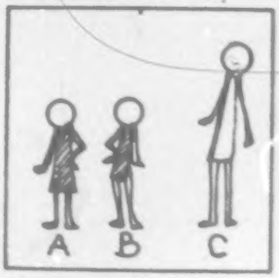
(2)



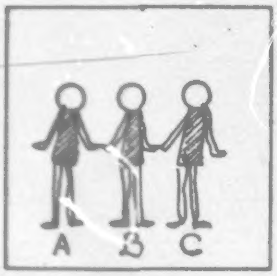
(3)



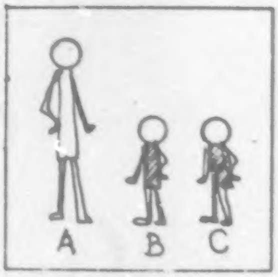
(4)



(5)



(6)



(7)

If A is taller than B, then B is taller than C.

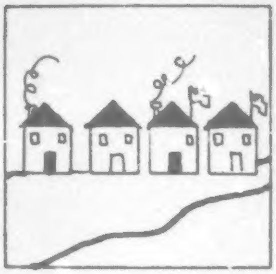
Use a procedure similar to the one described on page 11.

S-30

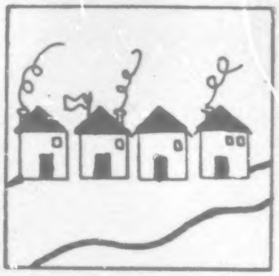
In each picture below you see a street with some houses. Read the conditional sentence at the bottom of this page and write the number of each picture which disagrees with it.

1,2,4 3

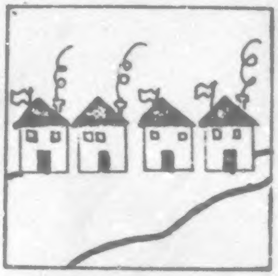
Students should check each picture to find out whether any house in it which has a chimney does not have a flag. One such house in a picture makes the picture disagree with the conditional sentence (or in other words: the sentence is false for that picture.



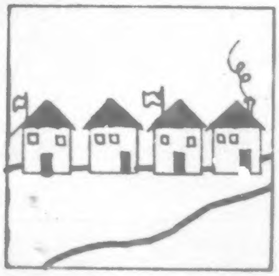
(1)



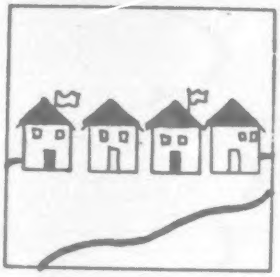
(2)



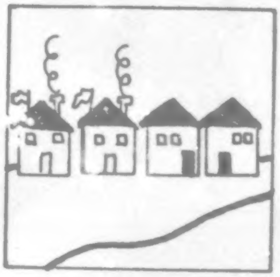
(3)



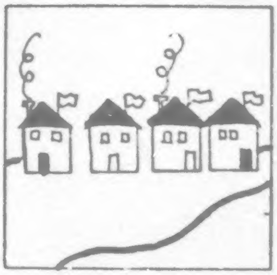
(4)



(5)



(6)



(7)

For all the houses in this street: If a house has a chimney, then it has a flag.

Each frame below contains nine numbers. Read the conditional sentence at the bottom of this page and write the number of each frame which disagrees with it.

2,4,5

For a given frame this sentence is true only if every even number in that frame is greater than 20. If in a given frame there is (at least) one even number which is less than, or equal to, 20, then this frame disagrees with the conditional sentence because not all even numbers in it are greater than 20. Odd numbers, of course, do not count.

21	22	23
24	25	26
27	28	29

(1)

15	16	17
18	19	20
21	22	23

(2)

13	15	17
19	21	23
25	27	29

(3)

12	14	16
18	20	22
24	26	28

(4)

11	12	13
14	15	16
17	18	19

(5)

13	15	17
19	22	24
26	28	30

(6)

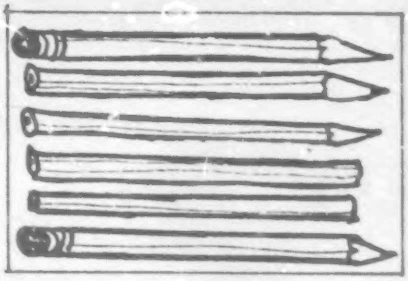
For all the numbers in this frame: If a number is even, then it is greater than 20.

S-32

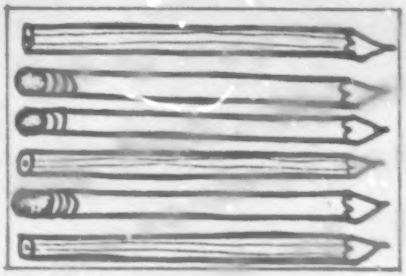
In each picture below you see a box of pencils. Some have an eraser, some don't. Some are sharp, some are not. Read the conditional sentence at the bottom of this page and write the number of each picture which disagrees with it.

4,5,7

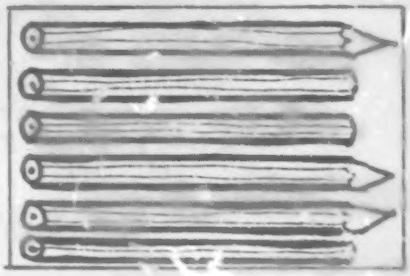
In each frame, one ought to look only at those pencils which have an eraser. Out of these one should look for those which are not sharp. If there is any pencil with an eraser but not sharpened in a given frame, then this frame disagrees with the sentence. (Pencils without an eraser may or may not be sharp.)



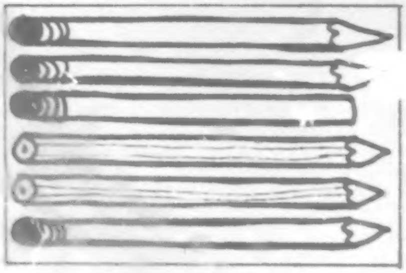
(1)



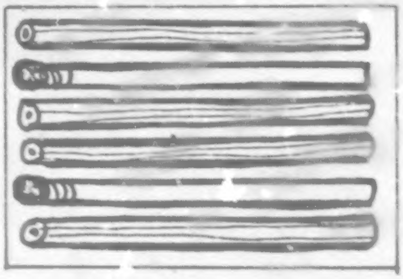
(2)



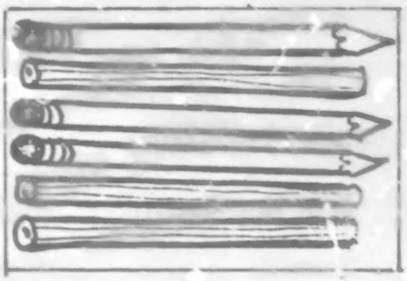
(3)



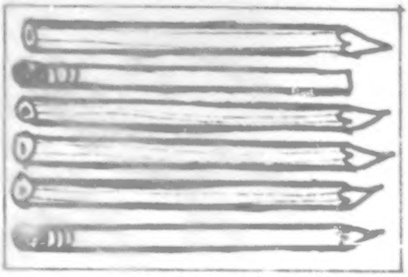
(4)



(5)



(6)



(7)

For all the pencils in this box: If a pencil has an eraser, then that pencil is sharp.

In each frame below there is a list of team members. Read the conditional sentence at the bottom of this page and write the number of each frame which disagrees with it.

2,3

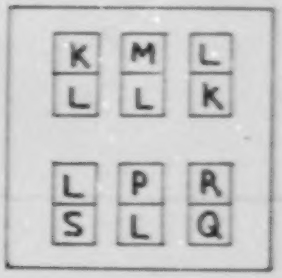
TEAM 1	TEAM 2	TEAM 3
MIKE	TOM	MARY
SCOTT	ARTHUR	NANCY
JERRY	DIANE	ANN
JOE	ALICE	RITA
(1)	(2)	(3)

This one may cause some difficulties. Students find pretty fast that frame 3 disagrees with the sentence. However, it takes some argument to persuade them that the sentence is false for frame 2. Try to lead those who have difficulties to explain the sentence using the word: must. For example - If a student belongs in this team (which one? - team 2) then this must be a boy. Does Diane belong in team 2? is this a boy? Well are all students in team 2 boys? Does it agree with the sentence?

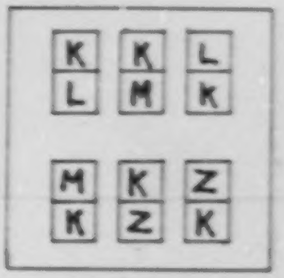
For all the students in the class: If a student belongs in this team, then that student is a boy.

In each frame below there are six cards. Each card has a letter on top, and a letter at the bottom. Read the conditional sentence at the bottom of this page and write the number of each frame which disagrees with it.

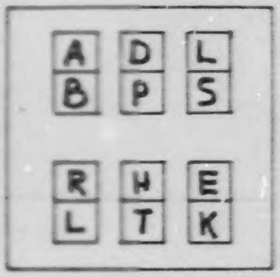
2,6


K
L
M
L
L
K
L
S
P
L
R
Q

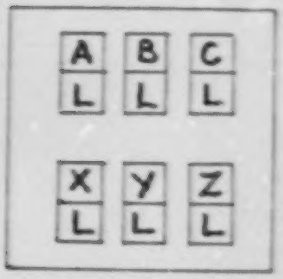
(1)


K
L
K
M
L
K
M
K
K
Z
Z
K

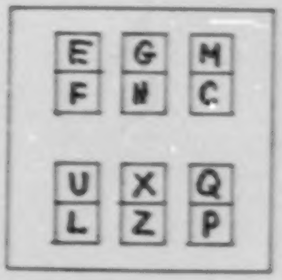
(2)


A
B
D
P
L
S
R
L
H
T
E
K

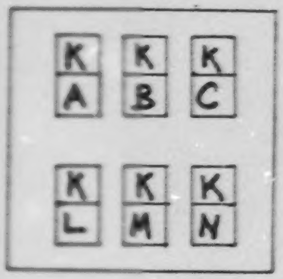
(3)


A
L
B
L
C
L
X
L
Y
L
Z
L

(4)


E
F
G
N
M
C
U
L
X
Z
Q
P

(5)


K
A
K
B
K
C
K
L
K
M
K
N

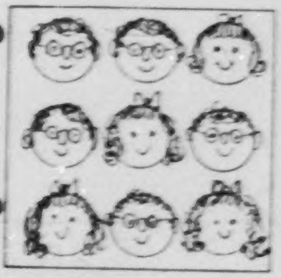
(6)

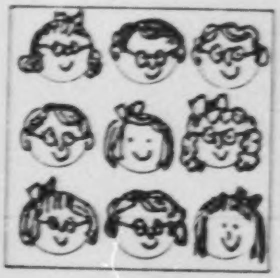
One card with K on top but a letter other than L at the bottom makes the whole frame disagree with the sentence, even if some other cards in that frame show K on top and L at the bottom. However, cards with L at the bottom do not *have to show K on top!*

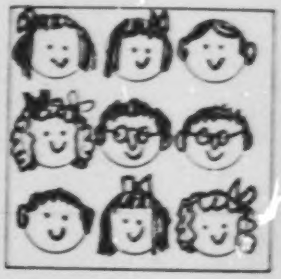
For all the two-letter cards in this frame: If there is a K on top of a card, then there is an L at the bottom of that card.

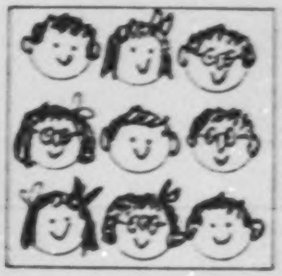
In each frame below there is a group of children. Read the conditional sentence at the bottom of this page and write the number of each frame which disagrees with it.

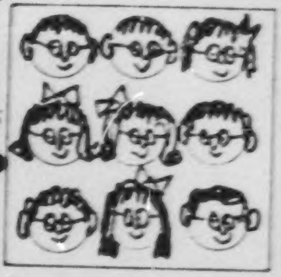
3,4,6,8,11,12

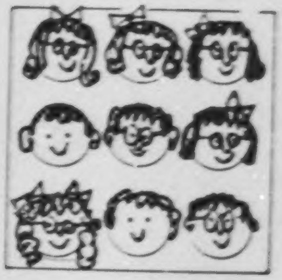

(1)

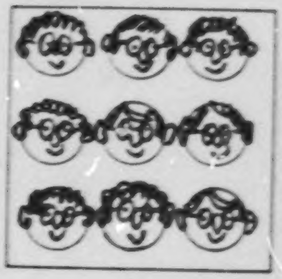

(2)

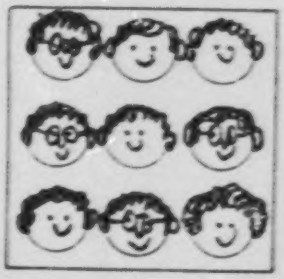

(3)

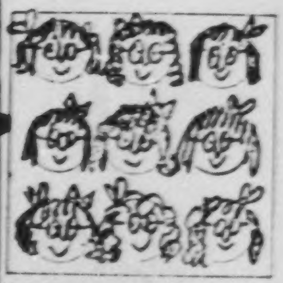

(4)

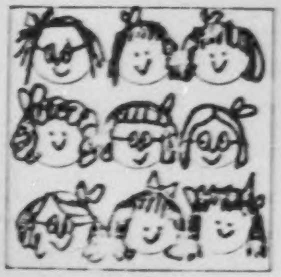

(5)

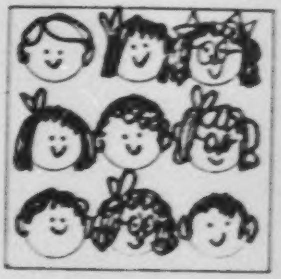

(6)

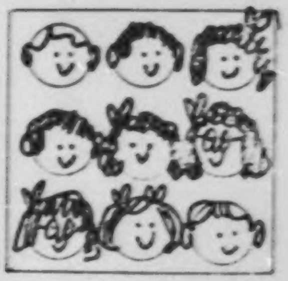

(7)


(8)


(9)


(10)


(11)


(12)

In this group of children: If a child does not wear a ribbon, then the child wears glasses. *In each frame we look for a child who does not wear a ribbon nor glasses.*

Numbers and Their Properties

Objectives: (1) "$(p \wedge q) \vee (\text{not-}p \wedge q) \vee (\text{not-}p \wedge \text{not-}q)$" implies "if p then q". "Not (p and not-q)" implies "if p then q".

(2) Using symbols (letters) for variables, constants, and relations (Introduction).

(3) Experience with MP, MT, AC, and DA.

Materials: A small hand blackboard (these activities are "chalk and talk," plus paper and pencil activities); or a magic slate

Duration: 2-3 sessions.

Administration:

Activity 1: Introduction of matrix representation of a conditional sentence.

This activity consists of three parts:

Part a takes about 5-10 minutes.

Part b takes about 15-20 minutes.

Part c takes about 15-20 minutes.

Part a. Make a chart on the blackboard like:

	> 35	not > 35
> 15		
not > 15		

T: Each of you think of a number, any number. Michael, what's your number?

Michael: 25.

T: Michael's number is 25. Is 25 greater than 15?

S: Yes.

T: (Point at the marginal title where it says > 15.) Well, it belongs in this row. Michael, is your number greater than 35?

Michael: No.

T: So, it belongs in this column (point at the marginal title saying: "not > 35"). So, I'll write 25, Michael's number, in here (put it in the upper right box).

T: Jane, what's your number? Where does it belong? Why? Jack, what's your number? Come write it down in the right place. Why did you put it there?

etc.

Part b.

T: (After each child has suggested his number.) Look, this box (point at the bottom left one) is still empty. Can any of you think of a number that belongs in this box? Think hard.

S: We can't; there is none; etc.

T: Why? What's the problem? We found numbers for this (upper right), and this (upper left), and this (lower right) box. Why can't we come up with a number for this one (lower left)?

S: Because there's no number that's greater than 35 and not greater than 15. (Or: because any number that's greater than 35 is also greater than 15; or: because a number that is not greater than 15 cannot be greater than 35; etc.)

T: That's right. If a number is greater than 35 (point at that marginal title and move your finger down along the column), then it must be greater than 15. It cannot be smaller than or equal to 15 (point at the lower left box) when it is greater than 35. (Write on the blackboard: If X > 35, then X > 15.) Also, if a number is not greater than 15 (point at that marginal title and move along its row), then it cannot be greater than 35. It must belong in here (point at the lower right box). Peter, can you repeat what I've said? Come here, show us how the chart tells us this. (The child will be encouraged to point at marginal titles and follow the rows and columns as he talks.

Part c.

T: In you papers in question 1 you have another chart. Fill it with the listed numbers.

(Children work on their own on question 1. When they finish let them work on 2, 3. Check their answers individually or make some of them tell their answers to the class. Leave the rest of the paper for next time.)

Activity 2. Puzzles

Part a.

T: If a number is > 35, then it is > 15. What sentence is that? (A conditional one.) How do you know? (It starts with "If.") Is that a true conditional sentence? (Yes.) Why? (Students will explain: no number can be greater than 35 without being greater than 15 because 35 is greater than 15, and so on.)

T: Here is a matrix like the one we had in your worksheets.

	> 15	not > 15
> 35	A	B
not > 35	C	D

(Draw it on the blackboard.)

I think of a number. I'll call my number X. My number X is greater than 35. (Write on the board X > 35.) Which box does it belong to?

S: The upper left, box A.

T: Why? O.K., come put X there. Now I think of another number. I'll call it Y. My number Y is greater than 15. (Write "Y > 15", under "X > 35".) Which box does it belong to?

S: Give us another clue. We can't tell yet. It may belong to either box of the left column, either A or C.

T: Come put Y in both. My new number is Z. Z is not > 35. (Write it under the previous two statements.) Where does Z belong?

S: You fooled us again: you didn't give us enough information. If Z is > 15 it belongs in C, the bottom left box; if it is not > 15, it belongs in D, the bottom right box (read ">" as "greater than").

T: Come put Z in both. I have another number now in my mind. I'll call it T. T is not > 15 (write it under the previous three statements). Can you tell which box T belongs in?

S: D.

T: Why? How come you don't need some more information?

S: T is not > 15, so it is certainly not > 35.

T: Does anybody have a number in mind? Andy, call your number a name. Tell us something about your number and we'll try to see if we can put it in our matrix, etc.

Part b.

Each student needs paper and pencil. Teacher needs a small blackboard on which he can write his secret numbers. Choose any true conditional sentence about numbers and their relations; for example, let's take the one the students worked on in their worksheets. Write it on the big blackboard in front of the class (preferably using symbols like: If X is > 60, then X is > 20.) Teacher will choose a number, will write it on the small blackboard, and hide it. Then the teacher will give the class a clue, which will be written on the big blackboard under the conditional sentence, and will ask a question about his hidden number. Students will be asked to answer, then the teacher will show his number.

All four questions will remain on the blackboard one next to the other (see examples, next page).

Say	Choose as your secret number	Write on the (big) blackboard
1. Jimmy, can you read this? What kind of a sentence is it. How do you know?		If $X > 60$, then $X > 20$
I choose a number. I'll not tell you what my number is but I can tell you it is greater than 60.	65 (hide it)	$X > 60$
Is my number greater than 20? (Yes)		Is $X > 20$?
	Show your hidden number	
2. Here is the same sentence.		If $X > 60$, then $X > 20$
I choose two numbers now! My new numbers both are greater than 20. Are they greater than 60? (NEC) Why? (Children give examples like: you may have chosen 22 and 23 or 22 and 62 or 62 and 63. In all cases X 20, but we can't tell whether X 60 or not.)	25 and 70 (hide either)	$X > 20$ Is $X > 60$?
	Show your numbers	
3. Here I choose another number.	30	If $X > 60$, then $X > 20$
My secret number this time is not greater than 60. Is it greater than 20? (NEC) Why?		X is not > 60 Is $X > 20$?
	Show it	
4. Last number I choose.	15	If $X > 60$, then $X > 20$
This time it is not greater than 20. Is it greater than 60? (No!) Why?		X is not > 20 Is $X > 60$?
	Show it	

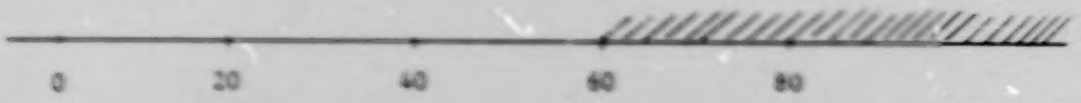
T: Now who wants to choose a number? Come here, Tom. Whisper it to me. Oh, Tom's number is greater than 20. Is it greater than 60?

(Or: Tom's number is greater than 60, is it greater than 20? Tom's number is not greater than 60, is it greater than 20? Or: Tom's number is not greater than 20, is it greater than 60?) Let different students whisper their number to you and you give the hint to the class. Children write their answers first, and then carry on the vote along with their reasoning. At the end, the child that chose the number says his number.

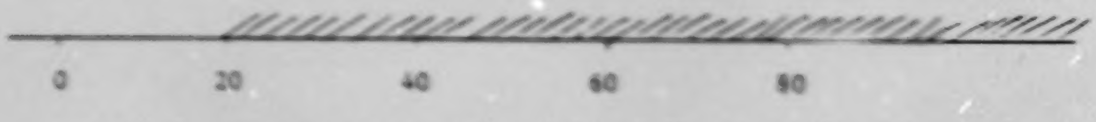
The following may help you phrase the questions:

If a child chooses a number	then the teacher says to the class
Greater than 35	His number is greater than 35, is it greater than 15? (Yes)
Between 15 and 35, or 35 itself (not 15!)	His number is not greater than 35. Is it greater than 15? (NEC) Or: his number is greater than 15, is it greater than 35? (NEC)
Less than 15, or 15	His number is not greater than 15. Is it greater than 35? (No)

Comment: If there's any student who has trouble answering these questions, a number line may help him visualize the problems. For example, in question 1 (above) the unknown number is somewhere here (shaded area):



so it is certainly to the right of 20, too. But in question 3 (above), the unknown number lies here (shaded area):



so it may or may not be to the right of 60. We can't tell. The given information is not enough to reach a decision.

Activity 3: Questions 4-15 in student's worksheets (see next pages for answer sheets).

Numbers and Their Properties

(1) Put each number listed below in the appropriate box in the matrix chart: (cross it off the list as soon as you put it in the matrix).

100, 45, 37, 29, 76, 12, 30, 1, 4, 42, 26, 15, 28, 35, 94,
56, 49, 13, 3, 24, 18, 6, 79, 52, 64, 19, 2, 91, 85, 22,
7, 33, 82, 77, 14, 34, 5, 11, 21, 36, 71, 84, 46, 51, 55.

	Greater than 20	Not greater than 20
Greater than 60	*100, 76, 94, 79, 64, 91, 85, 82, 77, 71, 84,*	
Not greater than 60	*45, 37, 29, 30, 42, 26, 28, 35, 56, 49, 24, 52, 22, 33, 34, 21, 36, 46, 51, 55,*	*12, 1, 4, 15, 13, 3, 18, 6, 19, 2, 7, 14, 5, 11,*

(2) In the chart you made in question (1), is there any empty box? *Yes*

Can you find a number to put in the empty box? *No*

Why? *Because there is no number that is at the same time greater than 60, and not greater than 20.*

(3) State the fact you discovered in question 2 as a conditional sentence. (Don't forget to start with the word: If).

a. If a number is greater than 60, then it is greater than 20.

b. If a number is not greater than 20, then it is not greater than 60.

In questions 4-7 read, think, answer, and give the reason

(4) If a number is greater than 60, then that number is greater than 20.

Tom's number is greater than 60.

Is Tom's number greater than 20? *Yes*

Why? *Because Tom's number is greater than 60 and 60 is greater than 20.*

(5) If a number is greater than 60, then that number is greater than 20.

Mary's number is not greater than 60.

Is Mary's number greater than 20? *NEC*

Why? *Mary's number is not greater than 60 but it may be greater than 20 (e.g., 30) or less than 20 (e.g., 10).*

(6) If a number is greater than 60, then that number is greater than 20.

Jill's number is greater than 20.

Is Jill's number greater than 60? *NEC*

Why? *Jill's number is not necessarily greater than 60. It may be greater than 60 (e.g., 70) or less than 60 (e.g., 40).*

(7) If a number is greater than 60, then it is greater than 20.

Jim's number is not greater than 20.

Is Jim's number greater than 60? *No*

Why? *If Jim's number was greater than 60 it would have been greater than 20, but it is not, so it can't be greater than 60.*

(8) The conditional sentence "If a number is greater than 60 then it is greater than 20" is a true statement about numbers and their relations. Invent some other true conditional sentences about numbers and their relations. Start each sentence in a new line. Don't forget: each conditional sentence should start with the right word, that is *If*.

Answers will vary. Here are some possible ones.

I. *If a number is greater than 100, then it is greater than 50.*

II. *If a number is less than 100, then it is less than 200.*

III. *If a number is divisible by 4, then it is divisible by 2.*

IV. *If a number has at least two digits, then it is greater than 9.*

(9) In questions 10-15, we'll use shortcuts:

a. X will stand for "my number".

b. The symbol < will stand for "less than".

c. Interpret (write in full words):

"X is < 5" means *My number is less than 5.*

"X is < 10" means *My number is less than 10.*

"If X is< 5, then X is <10" means: *If my number is less than 5, then my number is less than 10.*

d. Write using the shortcuts: My number is less than 30 $X < 30$

e. Write using the shortcuts: My number is less than 40 $X < 40$

f. Write using the shortcuts: If my number is less than 30, then my number is less than 40 If $X < 30$, then $X < 40$.

g. Interpret (in full words):

X is not < 10 My number is not less than 10.

Is $X < 10$? Is my number less than 10?

(10) Put each number listed below in the appropriate box in the matrix chart below. As soon as you do it, cross that number off the list.

100, 45, 37, 29, 76, 12, 30, 1, 4, 42, 26, 15, 28, 35, 94, 56, 49, 13, 3, 24, 18, 6, 79, 52, 64, 19, 2, 91, 85, 22, 7, 33, 82, 77, 14, 34, 5, 11, 21, 36, 71, 84, 46, 51, 55.

	< 30	not < 30
< 40	29, 12, 1, 4, 26, 15, 28, 13, 3, 24, 18, 6, 19, 2, 22, 7, 14, 5, 11, 21,	37, 30, 35, 33, 34, 36
not < 40		100, 45, 76, 42, 94, 56, 49, 79, 52, 64, 91, 85, 77, 82, 71, 84, 46, 51, 55

(11) In the chart you made in question (10), is there any empty box? *Yes*

Can you find a number to put in the empty box? *No*

Why? *Because every number that is less than 30, is less than 40 for sure*

(12) Tom: "If X < 30, then X < 40." (Read in full words: if my number is ...etc.)

Jill: "X < 30?" (Read: my number is ... etc.)

"Is X <40? "

Tom answered: *Yes, for sure.*

Why? *X < 30 and 30 < 40 therefore X < 40.*

(13) Tom: "If X < 30, then X < 40." (Read in full words!)

Jack: "X< 40."

"Is X < 30?"

Tom answered: *Can't tell. (NEC)*

Why? *Jack's number can be less than 40 and less than 30 (e.g., 20) or it can be less than 40 but not less than 30 (e.g., 35)*

(14) Tom: "If X < 30, then X < 40."

John: "X is not < 40."

"Is X < 30?"

Tom answered: *No*

Why? *John's number is not <40 so it can't be less than 30, because if it was, it would be less than 40 too.*

(15) Tom: "If X < 30, then X < 40."

Jane: "X is not < 30."

"Is X < 40?"

Tom answered: *NEC*

Why? *Jane's number is not less than 30, but it can still be either less than 40 (e.g., 35) or not (e.g., 60).*

(16) Invent some more conditional sentences about numbers and their relations. Start each sentence as a new line. Don't forget to begin every sentence with the right word, that is *If*.

I. *If a number is divisible by 3, then it's not a prime number.*

II. *If the ones digit of a number is 5, then this number is divisible by 5.*

III. *If a number is divisible by 6, then it is even.*

IV. *etc.*

Playing Cards

Objectives: a) not (p and not-q) implies: if p, then q.

b) $p \rightarrow q$ is equivalent to not-q $\rightarrow$ not p.

c) $p \rightarrow q$ and $q \rightarrow p$ are logically independent.

Materials: A deck of cards for each team (5-6 students). Exclude the jokers.

8 2x2 matrices as described below.

Slotted board (for Activity 1).

Duration: 2-3 sessions.

Administration:

Activity 1. Get acquainted with the cards.

a) In the slotted board put some hearts, some diamonds, some spades, and some clubs, each pattern in a separate row.

T: What is common to all the cards in the top row? In the bottom row? At the one right under the top? At the one right above the bottom?

Students will learn (or recall) the names for the different patterns.

b) Put cards on the slotted board. As teacher points at a card, students are to say: This is ten of clubs; this is king of hearts, etc.

c) The class will be divided into groups of 8-10 students. Each group member chooses a number from 1 to 10. One student in each group will be asked to distribute the cards among his peers such that each student gets all the cards with the number he chose and nothing but those cards. All the leftover cards are kept at the center of the table face down. Students' task is to describe the set of cards they hold. (Each student has four cards of the same number, 2 of them are black, 2 red, one card of each pattern: spade, club, diamond, heart). Let the students look at each other's sets and discover what is common to all sets. (Teacher will call students to show cards.)

T: Raise your card which shows a heart; raise your card which shows a diamond; etc. Show your red cards; show your black cards. Students with 5 of spades, raise your cards, etc.

Activity 2. Discovery of Conditional Sentences

Each team gets a 2x2 chart (matrices, see suggestions page 79 below). Each team puts a deck of cards (without the jokers) face down at the center of their table.

The task: You'll take turns. Each student will take one card from the pile and put it at the right place in the chart saying out loud, for example: "It is a heart and it is not red, so it belongs in here." When you finish putting all the cards in the chart, you will discover something. Write your discovery down.

When teams finish, the teacher asks for discoveries (e.g., there is no black-heart; there is no red club; all the red cards are either hearts or diamonds; there are 26 of each color, etc.) Encourage rephrasing of discoveries as conditional sentences (by reminding the students of the numbers chart they worked on previously). Right answers for each chart are listed below. The list is arranged in an increasing order of difficulty.

It is advised to have a felt or magnetic board and one deck of cards with felt or small magnets attached to each card for demonstration.

Charts for Activity 2 (R=red, B=black)

<table>
<tr><th>2x2 Matrix</th><th>Discovery</th><th>Conditional Sentences</th></tr>
<tr><td><table><tr><th></th><th>R</th><th>not-R</th></tr><tr><td>◇</td><td></td><td></td></tr><tr><td>not-◇</td><td></td><td></td></tr></table></td><td>No black diamonds
All diamonds are red</td><td>1. If a card shows ◇, then it is red.
2. If a card is not red, then it is not-◇</td></tr>
<tr><td><table><tr><th></th><th>B</th><th>not-B</th></tr><tr><td>♤</td><td></td><td></td></tr><tr><td>not-♤</td><td></td><td></td></tr></table></td><td>No red spade
All spades are black</td><td>1. If a card shows ♤, then it's black.
2. If a card is not black, then it does not show a spade.</td></tr>
<tr><td><table><tr><th></th><th>B</th><th>not-B</th></tr><tr><td>♧</td><td></td><td></td></tr><tr><td>not-♧</td><td></td><td></td></tr></table></td><td>No red club
All clubs are black</td><td>1. If a card is ♧, then it is black.
2. If a card is not black, then it is not-♧</td></tr>
</table>

	R	not-R
♡		
not-♡		

No black hearts

All hearts are red

1. If a card is ♡, then it is red.
2. If a card is not red, then it is not-♡

	B	not-B
◊		
not-◊		

No black diamonds

1. If a card is ◊ , then it is not black.
2. If a card is black, then it is not-◊

	R	not-R
♧		
not-♧		

No red clubs

1. If a card is ♧, then it is not red.
2. If a card is red, then it is not-♧

	B	not-B
♡		
not-♡		

No black hearts

1. If a card is ♡, then it is not black.
2. If a card is black, then it is not-♡

	R	not-R
♤		
not-♤		

No red spades

1. If a card is ♤, then it is not red.
2. If a card is red, then it is not-♤

Note: It is suggested that the teacher will teach the students how to rephrase their discoveries by pointing to a marginal title and following the two fields below it, or to its right, to see whether any of them is empty. For example, in the first chart the shaded area will stay empty because there are no diamonds which are not red. If it is a diamond (point to the upper left margin), then (pass your hand to the right) it must be red. We have no card in here (point to the shaded area). On the other hand, if a card is not red (point to the top right margin), then (pass your hand down) it cannot be a diamond. However, if a card is red (point to the top left margin), then (follow the chart down) it can be either a diamond or not, and if it is not a diamond, it can still be either red or not red.

	R	not-R
◇		
not- ◇		

Activity 3. Students need paper and pencil. Teacher writes on the blackboard 4 times in a row, a true conditional sentence discovered in Activity 2, say:

If ♡, then R	If ♡, then R	If ♡, then R	If ♡, then R

T: I'll ask you a question. You'll write your answers. (Hold a deck of cards. Pull out one that's a heart. Don't show its face.) I have a card in here. It's a heart. Is it red? (Write: ♡, under the left most sentence.) Teams' captains, count the answers and lead a discussion so that your team will agree unanimously on one answer. Team 1: What's your answer? Team 2: etc. (Ask for reasons!)

T: (Pull another card that's not a heart, preferably a diamond.) My card this time is not a heart, Is it red? (Write under the second conditional sentence: not- ♡; follow the procedure described above.)

T: (Pull another card, one that's red. Could be a heart this time.) My card now is red. (Write under the third conditional sentence: R.) Is it a heart? (Same procedure; encourage reasoning.)

T: (Pull another card, a black one.) Last time - my card is not red. Write not-R under the fourth conditional sentence. Is it a heart? (Same answering procedure). Who wants to pull a card? Come here, Denise. Don't show them your card. Show it to me. Oh, Denise's card is red. Is it a heart? (NEC) Or: Denise's card is not red, is it a heart? (No) Or: Denise's card is ♡. Is it red? (Yes) Or: Denise's card is not ♡, is it red? (NEC)

Repeat with other children.

Repeat this activity with another true conditional sentence for the deck of cards. Put each question under the one that matches it in logical type. Students who feel confident may be called to play the teacher's role in conducting the class.

Activity 4. Questions 1-4 in student's book may be used after Activity 2 or for summary. Questions 5-8 may be worked on another day, after completion of Activity 3.

Playing Cards

1. Write T for true or F for false next to each of the following sentences.
   - a. If a card shows ♡, then it's red. T
   - b. If a card is red, then it shows ♡. F
   - c. If a card is not red, then it doesn't show ♡. T
   - d. If a card does not show ♡, then it's not red. F

2. Write T for true or F for false next to each of the following sentences.
   - a. If a card shows ♤, then it's black. T
   - b. If a card is black, then it shows ♤. F
   - c. If a card is not black, then it doesn't show ♤. T
   - d. If a card doesn't show ♤, then it's not black. F

3. Write T for true or F for false next to each of the following sentences.
   - a. If a card shows ◇, then it is not black. T
   - b. If a card is not black, then it shows ◇. F
   - c. If a card is black, then it doesn't show ◇. T
   - d. If a card doesn't show ◇, then it's black. F

4. Complete the sentences.
   - a. If a card shows ♧, then it is not- red. T
   - b. If a card is red, then it doesn't show ♧. T
   - c. If a card is not- black, then it shows ♧. F
   - d. If a card shows ♧, then it is not black. F

The above questions should lead to an intuitive feeling of the generalizations:

1. *The truth of a conditional sentence does not imply the truth of its "flipped over" one (its converse).*
2. *A conditional sentence and its contrapositive are either both true or both false.*

In questions 5-8, read, think and answer.

5. **If a card shows ◇ , then it's not black.**

Barry's card shows ◇ .

Is Barry's card black? *No*

Why? *Because Barry's card shows and all diamond cards are not black.*

6. **If a card shows ◇ , then it's not black.**

Beth's card doesn't show ◇ .

Is Beth's card black? *NEC*

Why? *If Beth's card shows (= not-) then it is black. If it shows (= not-) then it is not black.*

7. **If a card shows ◇ , then it's not black.**

Benny's card is not black.

Does Benny's card show ◇ ? *NEC*

Why? *Cards that are not black are either hearts or diamonds, so Benny's can show a diamond but it does not have to show a diamond.*

8. **If a card shows ◇ , then it's not black.**

Brenda's card is black.

Does it show ◇ ? *No*

Why? *There is no black diamond. If it was a diamond, then it would be red, but it's black.*

Reasons are as important as much as right answers. Insist on getting reasons for any question.

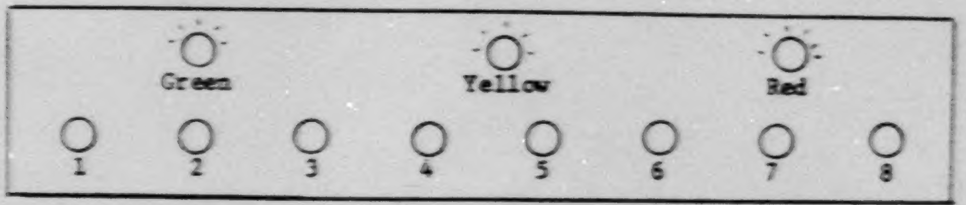
The colored light switch box

Objectives: (1) The structure of a conditional sentence.

(2) Use symbols (shortcuts) for sentences.

(3) Apply MP, AC, DA, AC and investigate their syntax.

Materials: The apparatus (six for each classroom).



The switch box consists of 8 switches and 3 colored bulbs. When someone operates it, he should remember to turn off any switch before turning on another one. (The ones with push-button switches are turned off automatically when you lift your hand.)

Duration of activity: 4-5 sessions.

Conducting the activities:

Activity 1: Get acquainted with the switchbox -(free play 10 min.)

Teacher will show the box; will push one or two switches to demonstrate how it is operated; will call one or two students to try and push switches; (will explain that these are expensive boxes and should be handled with care) and will give each group, of 5-7 students, a box.

Students will play freely with the box. Each student should be given a chance to touch and operate the box. (They may open it and look inside, they may push two or more switches at the same time--anything they want provided that it is done with care. THE BOX IS BREAKABLE.)

Activity 2: Get acquainted with the box(directed investigation 20-25 min.)

While the students are playing, as described above, the teacher will copy the following questions on the blackboard and distribute paper sheets for the students to write their answers.

Questions:

1. How many switches turn on the red light? Which ones do?
2. How many switches turn on the green light? Which ones?
3. How many switches turn on the yellow light? Which ones?
4. How many switches do not turn on the red light? Which ones?
5. How many switches do not turn on the green light? Which ones?
6. How many switches do not turn on the yellow light? Which ones?
7. How many switches turn on both green and yellow lights? Which ones?
8. How many switches turn on both green and red lights? Which ones do?
9. How many switches turn on both red and yellow lights? Which ones do?
10. Make a table. Put a plus sign (+) for each switch under each color that switch turns on. Put a minus sign (-) under each color that switch doesn't turn on. Can you find a pattern?

	Y	G	R
Switch 1	+	-	-
Switch 2			
3			
4			
5			
6			
7			
8			

Answers

	Y	G	R
Switch 1	+	-	-
Switch 2	+	-	+
3	+	+	-
4	+	+	+
5	-	-	-
6	-	-	+
7	-	+	-
8	-	+	+

Note: The purpose these questions serve is twofold: to release the tension and curiosity the box may raise on one hand, and to make children see the truth of a sentence like: "Switch #4 turns on the red light," where in fact this switch turns on the green light and the yellow light, too. In other words, it leads to the distinction between the sentence: "Switch #6 turns on only the red light," which is a false sentence, and "Switch #6 turns on the red light," which is a true sentence.

Activity 3: Generate some conditional sentences and test truth of others (30-35 min.)

Each team will have a switch box. Teacher needs one, too. Teacher will operate the switch box in front of the class.

T: (Push switch #1.) You see: If I push switch #1, then the yellow light comes on. Check your boxes. See if that's true. (Children verify.) Alright. So we have a true statement (write it in full words on the blackboard): If switch #1 is pushed, then the yellow light comes on. What kind of a sentence is it ?

S: This is a conditional sentence. It starts with "If."

T: Here is another conditional sentence. (Write it on the blackboard, under the first one, in full words. Ask a child to read it and verify its truth.): If switch #2 is pushed, then the yellow light is on.

Note: There will probably be some children who will say: "but the red light is on too." That's fine. We can have many true sentences about one switch. The main thing is that the teacher's sentence is one of the many possible true ones. Another one will be: "If switch #2 is pushed, then the yellow and red lights are on"; "If switch #2 is pushed, then the green light is not on"; "If "Switch #2 is pushed, then the red light comes on."

T: Who can come up with a conditional sentence about switch #3?

S: (Answers will vary. Teacher will list the sentences one under the other. Gradually ask for shortcuts--first for the colors. G.L.O. will stand for: "the green light is on"; R.L. not-o will stand for: "red light is not on," etc. Later on: S_1 will stand for "switch #1 is pushed." So "If S_7, then R.L.O." will be a shortcut for "If switch #8 is pushed, then the red light is on."

The following is a sample of true conditional sentences about the box.
1. If switch #1 is pushed, then the yellow light is on.
2. If switch #1 is pushed, then the green light is not on.
3. If switch #2 is pushed, then Y.L.O.
4. If switch #2 is pushed, then G.L. not-O.
5. If S_2, then R.L.O.
6. If S_3, then G.L.O.
7. If S_3, then R.L. not-O.
8. If S_4, then R.L.O.
9. $S_4 \rightarrow$ G.L.O.
10. $S_5 \rightarrow$ not Y.L.O.

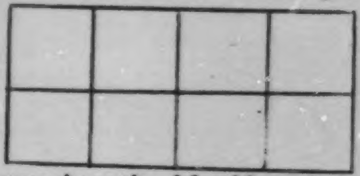
etc.

Activity 4: Logical puzzles

Pick a true conditional sentence about the switch box. Say:

$S_8 \rightarrow$ R.L.O.

Teacher only holds a switch box. Students need paper. Ask them to fold the paper in 4 parts, then 8 parts to get this form:



(Teacher will organize the blackboard in a way similar to the students' papers.)

T: I'll ask you 4 questions about this conditional sentence, $S_8 \rightarrow$ R. Let's write it 4 times, once in each of the 4 upper squares. Copy the question as I write it using shortcuts, and answer it.

Question 1. (Cover the lights only, push switch #8.)

Say	Write on the blackboard
	$S_8 \rightarrow$ R.L.O.
I pushed switch #8	S_8
Did the red light come on?	R.L.O.?

(Right answer: Yes)

Question 2. (Cover lights and switches. Push switch #6 or any other switch that turns on the red, but not switch #8.)

Say	Write on the blackboard
	S_8 R.L.O.
I pushed a switch. It's not switch #8	not S_8
Is the red light on?	R.L.O.?

(Right answer: NEC. During the discussion, show that some switches do and some don't turn on the red light, by pushing all switches one at a time.)

Question 3. (Cover lights and switches. Push switch #6 or any other switch that turns on the red light, e.g., 2, 4)

Say	Write on the blackboard
	S_8 R
I pushed a switch. The red light came on. (Show only the red light now)	R
Did I push switch #8?	S_8?

(Right answer: NEC. Discuss: It could be switch #8, but it does not have to be switch #8. Many other switches turn on the red light. Show it by pushing switches 2, 4, 6.)

Question 4. (Cover switches and lights. Push switch #1 or any other switch which does not turn on the red light, e.g., 3, 5, 7.)

Say	Write on the blackboard
	S_8 R
I pushed a switch. The red light did not come on. (Show the red bulb only.)	not R

Say	Write on the Blackboard
Did I push switch #8?	S_8?

(Right answer: No, because switch #8 turns on the red light, but the red light isn't on.)

Repeat the same thing again and again with different conditonal sentences. Make 4 questions for each conditonal sentence. After 2 or 3 sets of 4 questions, make sure to write them in the same order. Ask the students to study them and to find similarities and differences among questions in one row (same conditional sentence) and among sentences in the same column (same logical type). Some bright students may be asked now to try and invent, in a similar way, a set of 4 questions for a conditional sentence they choose.

Activity 5: Paper-pencil individual work with a switch box on each team table.

This work can be divided into three separate parts:

I. Questions 1-10

II. Questions 11-20.

III. Questions 21-25

Part I can be solved after Activity 3; Part II, after Activity 4. Part III: it is suggested that this part will be solved by the more successful students only.

The Colored Light Switch Box

The switch board consists of 8 switches and 3 colored bulbs. When you operate it remember to turn off any switch before turning on another one.

(1) Experiment:

In each line write just one color that makes the sentence true.

Example: If I push switch #4, then the red light is on.

Your turn:

(i): If I push switch #1, then the yellow light is on.
(ii): If I push switch #2, then the red light is on.
(iii): If I push switch #3, then the green light is on.

(2) (i): Nancy wrote: If I push switch #4, then the yellow light is on. Was Nancy right? yes

(ii): Complete the following sentence with one color to make it a true sentence, different than Nancy's sentence above: If I push switch #4, then the green light is on.

(3) Write 3 different colors to complete the following sentences so as to make each of them true (one color in each sentence):

(i): If I push switch #8, then the green light is on.
(ii): If I push switch #8, then the red light is on.
(iii): If I push switch #8, then the yellow light is not on.

(4) Let's write: S_1 as a shortcut for "Switch #1 is pushed";
S_2 as a shortcut for "Switch #2 is pushed."

Complete:

(i) We'll write S_4 as a shortcut for "switch #4 is pushed."

(ii) We'll write S_6 as a shortcut for "switch #6 is pushed."

(iii) We'll write S_7 as a shortcut for *switch 7 is pushed*.

(iv) We'll write S_5 as a shortcut for *switch 5 is pushed*.

(5) Experiment: Find at least three switches that turn the green light on, and use the shortcuts to write that they do it:

#3, *#4*, *#7*, *#8*.

(6) Let's use G.L.O. as a secret code for "The green light is on."

(a) Choose a secret code for "The red light is on." *R.L.O.*.

(b) Choose a secret code for "The yellow light is on." *Y.L.O.*.

(c) Interpret the following sentence into a full word sentence:

If S_2, then R.L.O.

If *switch #2 is pushed, then the red light is on.*

(7) Read, think, and answer. (Work slowly and carefully.)

Clues: (a) If S_2, then R.L.O. (Read: If switch #2 is pushed, then the red light is on.)

(b) S_2 (Read: Switch #2 is pushed.)

Question: Is the red light on? *yes*

Why? *Because switch #2 was pushed, and (a) says that in this case the red light must come on.*

(8) Read, think, and answer. (Slowly and carefully.)

Clues: (a) If S_8, then R.L.O. (In full words: *If switch 8 is pushed, then the red light is on.*)

b) S_8 (In full words: *Switch 8 was pushed*)

Question: Is the red light on? *Yes*

Why? *Similar to question 7.*

(9) Read, think, and answer.

Clues: (a) If S_8, then R.L.O. (Read: If switch #8 is pushed, then the red light is on.)

(b) R.L.O. (Read: the red light is on.)

Question: Was switch #8 pushed? *NEC*

Why? *Other switches turn the red light on, too. Switch also does but it is not the only one.*

(10) Read, think, and answer.

Clues: (a) If S_5, then Y.L.O. (In full words: *If switch 5 is pushed, then the yellow light is on.*)

(b) Y.L.O. (*The yellos light is on.*)

Question: Was switch #5 pushed? *NEC*

Why? *The yellow light may be turned on by some other switches, too. There is no way to tell whether #5 was or was not pushed.*

(11) Sentences which start with the word "If" are called "conditional sentences," because they always tell us something about conditions. You hear and use conditional sentences many times, every day. Invent two conditional sentences (be imaginative!).

If *(Answers will vary. Reinforce ones that make sense and even more so - ones for which the converse does not*

If *make sense.)*

(12) There is another word besides "If" which usually appears in conditional sentences.

Have you discovered it? Write it here: *then*

Notice: The word "If" always appears in any conditional sentence, the word "then" is sometimes omitted.

In questions (13), (14) read both clues out loud (in full English) and answer the question.

We'll use: "G.L.O.?" as a shortcut for "Is the green light on?"

We'll use: "S_1?" as a shortcut for: "Is switch #1 pushed?"

(13) Read (in full words), think, and answer.

Clues: (a) If S_1, then Y.

(b) S_1 .

Question: Y. L.O. *Yes*

because *(Similar to 7.)*

(14) Read (in full words), think, and answer.

Clues: (a) If S_1, then Y

(b) Y .L.O.

Question: S_1? *NEC*

Explain why *(Similar to 9.)*

(15) (I) Puzzle 13 is somehow similar to two puzzles you have done before in this paper. Go back and try to find which ones. *7, 8*

(II) Puzzle 14 also may remind you of two puzzles you have already worked out in this paper. Which ones? *9, 10*

Note: The right answers to I and II are different! Check your answers again.

(16) We'll write not-S_1 as a shortcut for: "Switch #1 was not pushed."

Complete:

How would you shorten the following?

(I) We'll write not S_2 as a shortcut for: "Switch #2 was not pushed."

(II) We'll write not S_7 as a shortcut for: "Switch #7 was not pushed."

(III) We'll write not S_5 as a shortcut for: "Switch #5 was not pushed."

(IV) We'll write not-S_1 as a shortcut for " Switch 1 was not pushed. ".

(V) We'll write not-S_4 as a shortcut for " Switch 4 was not pushed. ".

(17) We have been using G.L.O. as a shortcut for "the green light is on."

Invent a shortcut for the opposite sentence: "the green light is not on." not-G.L.O.

What would you write for: "The yellow light is not on?" Not Y.L.O.

(18) Interpret (write in full words) the sentence:

If not-Y.L.O., then not S_1.

If the yellow light is not on, then switch 1 was not pushed.

Is it a true sentence? Yes

Why? Because switch 1 always turns on the yellow light.

In questions (19), (20) read both clues out loud (in full English) and see if you can draw any conclusion. If there is no conclusion that can be drawn, write that fact down and explain why.

(19) Read (in full words), think, and answer.

Clues: (a) If S_1, then Y.L.O.

(b) Not-S_1

Question: Y.L.O.? *NEC*

Why? *The yellow light could be turned on by another switch.*

(20) Read (in full words), think, and answer.

Clues: (a) If S_1, then Y.L.O.

(b) Not-Y.L.O.

Question: S_1? *No* Why? *Because switch 1 always turns on the yellow light.*

In questions (21)-(24) we'll help you to invent some questions. Answer every question you invent.

(21) Invent a question which looks very much like question (7) but starts as follows:

Clues: (a) If switch #8 is pushed, then the red light is on.

(b) *Switch #8 was pushed.*

Question: *Did the red light come on?*

Answer: *Yes* Why? *That's what (a) says.*

(22) Invent clue (b) to get a question which looks very like question (9) but starts as follows:

Clues: (a) If switch #8 is pushed, then the red light is on.

(b) *The red light is on.*

Question: *Was switch 8 pushed.*

Answer: *NEC* Why? *Red light is turned on not only by switch 8.*

(23) Invent clue (b) and a question which looks very much like question (19) but starts as follows:

Clues: (a) If switch #8 is pushed, then the red light is on.

(b) *Switch 8 was not pushed.*

Question: *Is the red light on?*

Answer: *NEC* Why? *There are other switches that may have turned on the red light.*

(24) Invent clue (b) and a question which looks very much like question (20) but starts as follows:

Clues: (a) If switch #8 is pushed, then the red light is on.

(b) *The red light is not on.*

Question: *Was switch 8 pushed?*

Answer: *No* Why? *Because if it was pushed, then the red light would light.*

(25) Invent another set of 4 questions like questions (21)-(24) starting with any conditional sentence you choose.

(I) Clues: (a) If S_4 then G.L.O.

(b) S_4

Question: G.L.O?

Answer: Yes Why? Similar to 21.

(II) Clues: (a) $S_4 \rightarrow$ G.L.O

(b) G.L.O.

Question: S_4?

Answer: NEC Why? Similar to 22.

(III) Clues: (a) $S_4 \rightarrow$ G.L.O

(b) not S_4

Question: G.L.O.?

Answer: NEC Why? Similar to 23.

(IV) Clues: (a) $S_4 \rightarrow$ G.L.O

(b) not-G.L.O.

Question: S_4?

Answer: No Why? Similar to 24.

Prepare a Quiz

Objective: To increase awareness of the syntactical differences among MP, MT, AC, and DA.

Method: This will be done in two steps:

(I) The general form of a conditional sentence. (See below.)

(II) Construction of 4 puzzles from one conditional sentence and preparation of pupil's made electric cards.

Materials: List of sentences. Materials for electric cards (wires, cardboard, stickers).

Duration: 4-5 sessions.

Procedure for Step (I): **Activity 1**: **Coding-Decoding Game**

The teacher will first discuss the structure of a conditional sentence. The notion of a conditional sentence is by now familiar to the students, and so the discussion may take a form of the following kind:

Teacher: We have been doing lots of work with conditional sentences lately. Can any of you remind us of some of the conditional sentences you had? (Teacher will write students' suggestions on the blackboard, one under the other.)

Student 1: If a card is ◇, then it is R.

Student 2: If a number is > 150, then it is > 100.

Student 3: If Ryan is out, then his mother is out.

Student 4: If it is raining, then Carol wears her boots.

Student 5: If S_6, then G.

T: Here are some sentences. (Post a list. See a sample on page 103.)

Now here is my conditional sentence (write on the blackboard and say:) "If (a), then (e)."

Can you decode it?

S: If it is dark, then it is scary.

T: If (g), then not (a). (Write it down under all the others.)

S: If the sun is up in the skies, then it is not dark.

T: Do any of you want to try to invent a conditional sentence?

S_1: If (p), then (s).

T: Who knows what this sentence says?

S_2: If it it snowing, then Fred drives slowly.

T: Very good. Anybody else want to invent a conditional sentence out of these?

Students will suggest their sentences, and write them one under the other. (Student may suggest conditional sentences including negations, e.g.: If (a), then not (f), which means: If it is dark, then Fred does not read a book. Students may also suggest ridiculous sentences like "If (m), then (g)...) When there are enough examples on the blackboard (or when students are getting tired of the coding-decoding game), teacher will lead to the generalization . Here is one of many possible ways:

T: That will be enough. Thank you. Now, all these are conditional sentences (write these words on top). Can you find anything that's common to all of them?

S_1: They all start with If (Teacher underlines all the "Ifs" on the blackboard.)

S_2: They all are divided into two parts by a comma. (Teacher circles the commas.)

S_3: They all have the word "then", kind of in the middle (Teacher will underline all the "thens".)

T: I want to invent a new conditional sentence. With what word shall I start?

S: With the word: If.

T: (Write under the examples on the blackboard: If.) O.K., I have If something (put dots...right after the word "If" all the way up to the commas in the above examples.) Now what?

S: Now, put a comma, write "then" and say something else.

T: Alright, here we are: If..., then.... This is the form of any conditional sentence.

Students will be asked to write down (copy from the blackboard) in their notebooks the following:

Conditional Sentences

Examples:

1. If a number is 150, then it is 100.

2. If Ryan is out, then his mother is out.

3. If S_6, then G.

In general:

If ...., then ....

Activity 2: Worksheets

T: Let's work on questions (1)-(3) in your book.

A Sample of Sentences for Construction of Conditional Sentences

(To be posted in front of the class)

a) It is dark.
b) Fred wears his shoes.
c) His feet will hurt.
d) He goes out.
e) It is scary.
f) Fred reads a book.
g) The sun is up in the skies.
h) It is raining.
i) He feels good.
j) It is vacation time.
k) He works hard.
l) The streets are slippery.
m) Fred wears his coat.
n) He sleeps late.
o) He needs some quiet around.
p) It is snowing.
q) He takes a shower.
r) It is hot.
s) Fred drives slowly.
t) Fred is tired.

Procedure for Step (II)

Activity 3: Electric Cards

Students will be asked to answer four questions (MP, MT, AC, and DA items) based upon one conditional sentence. They will get electric cards in sets of four cards with the same conditional sentence on all four of them. After answering them they will be called upon to study the four questions to find differences and commonalities among them.

Activity 4: Coding-Decoding Game

Using the above sample of sentences (Activity 2), the teacher will write four questions on the blackboard using shortcuts (one to be answered at a time, needless to say).

For example:

If (j), then (i) (j) (i)?	If (j), then (i) not (j) (i)?	If (j), then (i) (i) (j)?	If (j), then (i) not (i) (j)?

Children will decode and answer each question.

The next four questions will be written down under these, each one under the one having the same logical type.

After a few examples, a student will be called upon to invent a puzzle based on a conditional sentence (using the list). The teacher will ask him to put it on the blackboard in the proper column. After one student's puzzle is answered, the class will be challenged to invent another puzzle with the same conditional sentence. (Each puzzle will be answered right after its invention, but it is not recommended to put the answers on the blackboard for this may push too fast toward the algorithm which may then lead to automatical work that is boring on one hand, and does not involve logical thinking on the other hand.) Students will be happy to lead the class for their coded puzzle.

Note: Each student is expected to discover the four possible relations between the conditional sentence (first clue) and the second clue which is either the antecedent or the consequent or any of their negations. They are not expected, however, to be able to express this discovery verbally. Their discovery will be expressed by their ability to construct all four kinds for a particular conditional sentence. The above activity can also be conducted as a group activity where students take turns presenting their coded puzzle.

Activity 5: Prepare a Quiz - Team Contest (4-5 students in each team)

1st Game

Each team member will invent and write four puzzles; possibly, but not necessarily, built up of one conditional sentence. (Note: Some students will need help in inventing a conditional sentence.) Each team member will write the answers to his puzzles in parentheses under the puzzle. Team's captain will then collect the puzzles, let the teacher check the answers, and exchange seats with the other team's captain. He will now read his team's puzzles to the competing team, which will discuss them and arrive at an answer. The team will score one point for any right answer. The team with the higher total score is the winner.

2nd Game

The teacher will announce an answer for which students will prepare puzzles. E.g., prepare only puzzles for which the right answer is: not enough clues (or: either yes or no, or: no, or: yes). Each team member invents and writes three puzzles. The teams exchange puzzles. The task now is to check the puzzles and to find those for which the answer is not the one announced by the teacher. The team's captain reads the other team's puzzles to his team members. Right answer will be discussed and the team will score a point for any puzzle they find for which the right answer is not the answer announced by the teacher. The captain will separate out those puzzles and will ask teacher's approval of the fact that the answer differs from the preassigned one. The team with the higher total score is the winner.

Note: This game will lead some students to discover that denying the antecedent and affirming the consequent always lead to the answer: not enough clues. It is doubted that they will be able to express it but their consistent behavior will indicate it.

Preparation for Activity 5 is given in the Electric Cards contest activity and in the Switch Board activity.

Activity 6: Prepare Electric Cards Project

Each child will follow the directions in question 4 of the worksheet. Step I can be taken right after Activity 2. Step II - Teacher should direct the students to choose a conditional sentence that makes sense and that is irreversible, namely it's"flipped over"sentence does not make sense. ($p \rightarrow q$ makes sense but $q \rightarrow p$ does not.)

Step III - Right after Activity 4.

Step IV and on - Final Project.

Prepare a Quiz

(1) **Here are two sentences:**

(a) **This is a chicken.**

(b) **It has two legs.**

Interpret (in full words) the sentence:

If (a), then (b).

If this is a chicken, then it has two legs.

(2) **For the following conditional sentence fill the blanks:**

$$\text{If } \underbrace{\text{she has a headache}}_{c}\text{, then } \underbrace{\text{she takes an aspirin.}}_{d}$$

(c) **is:** *She has a headache.*

(d) **is:** *She takes an asprin.*

Interpret (in full words) the sentence:

If (d), then (c).

If she takes an asprin, then she has a headache. (This sentence does not follow from the previous one!)

(3) **Invent a sentence (f) and write it down in the blank:**

(e) **This is Ronald.**

(f) *(answers will vary, e.g.:) Mary'll be happy.*

Interpret (in full words) the following sentence:

If (e), then (f).

If this is Ronald, then Mary will be happy.

(4) Project: Prepare your own electric cards by following the directions below:

Step I. Invent a conditional sentence of your own. (Remember to start with the right word.)

(Answers will vary.) If he does not pass the driving test, then he will not buy a car.

Call (g) the part which comes right after the word "If" of your sentence.

(g) reads *He does not pass the driving test.*

Call (h) the second part of your sentence. This part comes right after the word "then."

(h) reads *He will not buy a car.*

Write your conditional sentence in symbols:

If g, then h. (or: g → h)

Step II. Have your sentence checked by your teacher. Change it if necessary. *(See teacher's manual for how to check.)*

Step III. Construct four different puzzles from the sentence you invented in (4) by interpreting (in full words) the symbols below.

Puzzle 1 *Answers will vary.*

clue a: g → h *If he does not pass the driving test, then he will not buy a car.*

clue b: g *He did not (*) pass the driving test.*

Question: *Will he buy a car?*

Answer: *No.*

Puzzle 2

clue a: g → h *If he does not pass the driving test, then he will not buy a car.*

clue b: not-g *He passed (*) the driving test.*

Question: *Will he buy a car?*

Answer: *NEC (He may be unable to afford it.)*

(*) Some changes in tenses may be necessary!

Puzzle 3

Clue a: g → h *If he does not pass the driving test, then he will not buy a car.*

clue b: h *He does not (*) buy a car.*

Question: *Did he pass the driving test?*

Answer: *NEC*

Puzzle 4

Clue a: g → h *If he does not pass the driving test, then he will not buy a car.*

clue b: not-h *He bought (*) a car.*

Question: *Did he (*) pass the driving test.*

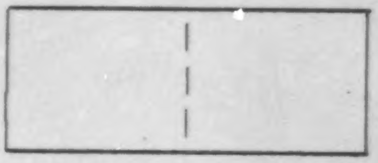
Answer: *Yes. (Otherwise he wouldn't buy a car!)*

(*) Some grammatical modifications may be needed.

Step IV. Have your puzzles checked by your teacher. Change them if necessary. *(Check tenses of verbs so that puzzles sound right.)*

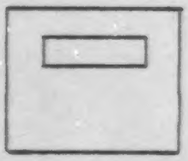
Step V. Get 4 stickers and copy each puzzle on a sticker.

Step VI. Get a cardboard and divide it into two equal parts using a pencil like this:



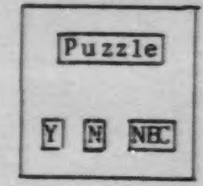
And fold it over like this:

Step VII. Peel a puzzle sticker and paste it on top.

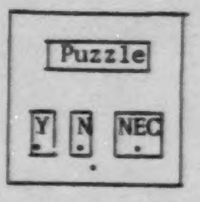


Step VIII. Repeat steps VI, VII, to get four cards.

Step IX. Get answer stickers and put them one next to the other on each card like this:



Step X. Get 4 paper fasteners for each card (16 altogether) and put them into the cards like this:



Make sure the fasteners don't touch each other at the back.

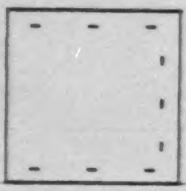
Step XI. Get a wire and wire the bottom fastener to the fastener which belongs to the right answer.

Make sure you do it right.

Make sure the wire does not touch other fasteners.

Step XII. Test your cards with the electric tester. Make sure the bulb lights on the right answer.

Step XIII. Staple your card like this:



Step XIV. Write your name on the back.

Step XV. Challenge a friend with your cards.

APPENDIX 7.2

FINAL VERSION OF THE TEST

Notice: The upper half of each page belongs to version T, and the bottom half belongs to version T'. Please refer to table 3.4 (page 108) for item's logical form and negation mode.

-1-

Version T (green)

Students name ______________________________
(First) (Last)

Birth-date ______________ Age ______ Boy or girl? ______
(Month) (Year)

Grade ______ Teacher's name ______________________

Today's date ______________ School ______________

-2-

Version T' (yellow)

Student's name ______________________________
(First) (Last)

Birth-date ______________ Age ______ Boy or girl? ______
(Month) (Year)

Grade ______ Teacher's name ______________________

Today's date ______________ School ______________

T - 2 -

(This page was read to the class before handing out the puzzle-book).

We are going to have a team-contest, in puzzle solving. Each student will work on his own puzzle book. The points each member gets will sum up to the team's score. We'll have a yellow team and a green team (show a sample puzzle book of each).

- You will get one point for each correct answer -

- You will lose one point for each wrong answer -

Therefore, each time you try a puzzle, try hard to choose the right answer so that you and your team will earn a point. But if you are really not sure, leave the puzzle unanswered, so that your team will not lose a point. Just go to the next puzzle. If you put a sign like * next to a puzzle you skip, then you can come back to it when you finish.

If you find a word that you do not know, raise your hand, and I will explain that word to you. THIS IS NOT A TEST OF YOUR READING, so ask about any word you are not sure of.

These puzzles are just for your fun, and you may take all the time you want on them. There is NO TIME LIMIT. The competition is *only* on the number of right answers you get. So, work carefully and help your team win.

When you finish, check your answers again, then you may draw some pictures or design on the back of any page (show where), but PLEASE, REMAIN SEATED and QUITE until I collect the papers. As you get your puzzle book, please fill in the blanks on the front page. (Distribute puzzle books.)

T' - 2 -

We are going to have a team-contest, in puzzle solving. Each student will work on his own puzzle book. The points each member gets will sum up to the team's score. We'll have a yellow team and a green team (show a sample puzzle book of each).

- You will get one point for each correct answer -

- You will lose one point for each wrong answer -

Therefore, each time you try a puzzle, try hard to choose the right answer so that you and your team will earn a point. But if you are really not sure, leave the puzzle unanswered, so that your team will not lose a point. Just go to the next puzzle. If you put a sign like * next to a puzzle you skip, then you can come back to it when you finish.

If you find a word that you do not know, raise your hand, and I will explain that word to you. THIS IS NOT A TEST OF YOUR READING, so ask about any word you are not sure of.

These puzzles are just for your fun, and you may take all the time you want on them. There is NO TIME LIMIT. The competition is *only* on the number of right answers you get. So, work carefully and help your team win.

When you finish, check your answers again, then you may draw some pictures or design on the back of any page (show where), but PLEASE, REMAIN SEATED and QUIET until I college the papers. As you get your puzzle book please fill in the blanks on the front page. (Distribute puzzle books.)

T - 3 -

(Please, follow my reading.)

The Puzzles

In each puzzle we give you two clues, then we ask a question. For each question there are three possible answers: ( ) Yes, ( ) No, ( ) Not Enough Clues. Decide which is the right answer for each puzzle, and indicate this by marking x in the parentheses to the left of that answer.

In some puzzles we did not give you enough clues to reach a yes or a no answer. So, if you think there are not enough clues in a given puzzle, be sure to put your x next to the answer - Not enough clues. In some puzzles, of course, there are enough clues to reach a yes or a no answer. In such cases, be sure to mark your x next to yes or no.

Remember: In each case MARK ONLY ONE ANSWER

Let's do two examples together:

T' - 3 -

(Please, follow my reading)

The Puzzles

In each puzzle we give you two clues, then we ask a question. For each question there are three possible answers: ( ) Yes, ( ) No, ( ) Not Enough Clues. Decide which is the right answer for each puzzle, and indicate this by marking x in the parentheses to the left of that answer.

In some puzzles we did not give you enough clues to reach a yes or a no answer. So, if you think there are not enough clues in a given puzzle, be sure to put your x next to the answer - Not enough clues. In some puzzles, of course, there are enough clues to reach a yes or a no answer. In such cases, be sure to mark your x next to yes or no.

Remember: In each case MARK ONLY ONE ANSWER

Let's do two examples together:

Example 1. Clues: (a) If it rains tomorrow, then it will be cloudy tomorrow.

(b) It will not be cloudy tomorrow.

Question: Will it rain tomorrow?

( ) Yes ( ) No ( ) Not enough clues

The answer is: No. It will certainly not rain tomorrow, because if it does, then, according to the first clue, it will be cloudy. But the second clue says it will not be cloudy. So, No is the right answer. Mark x in the parentheses to the left of that answer.

T' - 4 -

Example 1 Clues: (a) If it rains tomorrow, then it will be cloudy tomorrow.

(b) It will not be cloudy tomorrow.

Question: Will it rain tomorrow?

( ) Yes ( ) No ( ) Not enough clues

The answer is: No. It will certainly not rain tomorrow, because if it does, then, according to the first clue, it will be cloudy. But the second clue says it will not be cloudy. So, No is the right answer. Mark x in the parentheses to the left of that answer.

T - 5 -

Example 2. Clues: (a) If it rains tomorrow, then it will be cloudy tomorrow.

(b) It will not rain tomorrow.

Question: Will it be cloudy tomorrow?

( ) Yes ( ) No ( ) Not enough clues

The right answer is: Not enough clues. The first clue says nothing about what will happen if it does not rain. It may be cloudy tomorrow even though it does not rain, or it may be not cloudy and not raining. There are not enough clues to definitely decide whether the answer is yes or no. Mark this answer.

Do you have any questions before we begin?

Now, turn the page and do your best.

T' - 5 -

Example 2. Clues: (a) If it rains tomorrow, then it will be cloudy tomorrow.

(b) It will not rain tomorrow.

Question: Will it be cloudy tomorrow?

( ) Yes ( ) No ( ) Not enough clues

The right answer is: Not enough clues. The first clue says nothing about what will happen if it does not rain. It may be cloudy tomorrow even though it does not rain, or it may be not cloudy and not raining. There are not enough clues to definitely decide whether the answer is yes or no. Mark this answer.

Do you have any questions before we begin?

Now, turn the page and do your best.

T

1. Clues: (a) If Mary is seven, then she is too young for that summer camp.

(b) Mary is seven.

Question: Is she too young for that summer camp?

( ) Yes ( ) No ( ) Not enough clues

T'

1. Clues: (a) If the wind blows from the west, then the clouds will go away.

(b) The wind blows from the west.

Question: Will the clouds go away?

( ) Yes ( ) No ( ) Not enough clues

T

2. Clues: (a) If there is a policeman at the corner, then Don waits on the sidewalk.

(b) Don waits on the sidewalk

Question: Is there a policeman at the corner?

( ) Yes ( ) No ( ) Not enough clues

T'

2. Clues: (a) If the aquarium is dirty, then the goldfish will die.

(b) The goldfish has died.

Question: Was the aquarium dirty?

( ) Yes ( ) No ( ) Not enough clues

T

3. Clues: (a) If the weather is not warm, then Cindy does not go swimming.

(b) The weather is not warm.

Question: Is Cindy going to swim?

( ) Yes ( ) No ( ) Not enough clues

T'

3. Clues: (a) If it is not a children's show, then Ronald does not watch it.

(b) This is not a children's show.

Question: Does Ronald watch it?

( ) Yes ( ) No ( ) Not enough clues

T

4. Clues: (a) If Terry has a fever, then he will not go to school tomorrow.

(b) Terry will not go to school tomorrow.

Question: Does Terry have a fever?

( ) Yes ( ) No ( ) Not enough clues

T'

4. Clues: (a) If someone plays too much football, then he does not do enough homework.

(b) Steve does not do enough homework.

Question: Does Steve play too much football?

( ) Yes ( ) No ( ) Not enough clues

T

5. Clues: (a) If Sue's desk is not cleaned up, then she has to stay after school.

(b) Sue does not have to stay after school.

Question: Is Sue's desk cleaned up?

( ) Yes ( ) No ( ) Not enough clues

T'

5. Clues: (a) If the wind does not change, then our sailboat will approach the dock.

(b) Our sailboat does not approach the dock.

Question: Did the wind change?

( ) Yes ( ) No ( ) Not enough clues

T

6. Clues: (a) If the dog out there has black spots, then it is not my dog.

(b) The dog out there does not have black spots.

Question: Is it my dog?

( ) Yes ( ) No ( ) Not enough clues

T'

6. Clues: (a) If a student takes Spanish, then he does not take French.

(b) Jeff does not take Spanish.

Question: Does he take French.

( ) Yes ( ) No ( ) Not enough clues

T

7. Clues: (a) If someone plays too much football, then he does not do enough homework.

(b) Steve does not do enough homework.

Question: Does Steve play too much football?

( ) Yes ( ) No ( ) Not enough clues

T'

7. Clues (a) If Terry has a fever, then he will not go to school tomorrow.

(b) Terry will not go to school tomorrow.

Question: Does Terry have a fever?

( ) Yes ( ) No ( ) Not enough clues

T

8. Clues: (a) If it is a holiday, then the library is not open.

(b) The library is open.

Question: Is it a holiday?

( ) Yes ( ) No ( ) Not enough clues

T'

8. Clues: (a) If the coat is black, then it is not Jill's.

(b) This is Jill's coat.

Question: Is it black?

( ) Yes ( ) No ( ) Not enough clues

T

9. Clues: (a) If the wind does not change, then our sailboat will approach the dock.

(b) Our sailboat does not approach the dock.

Question: Did the wind change?

( ) Yes ( ) No ( ) Not enough clues

T'

9. Clues: (a) If Sue's desk is not cleaned up, then she has to stay after school.

(b) Sue does not have to stay after school.

Question: Is Sue's desk cleaned up?

( ) Yes ( ) No ( ) Not enough clues

T

10. Clues: (a) If father fills up his car with gas, then he cleans the windshield.

(b) Father does not fill up his car with gas.

Question: Is he cleaning the windshield?

( ) Yes ( ) No ( ) Not enough clues

T'

10. Clues: (a) If it is Monday, then mother stays at home.

(b) It is not Monday.

Question: Is mother staying at home?

( ) Yes ( ) No ( ) Not enough clues

T

11. Clues: (a) If Jack is not in the race, then his team will win.

(b) Jack is not in the race.

Question: Will his team win?

( ) Yes ( ) No ( ) Not enough clues

T'

11. Clues: (a) If Janet does not come home in time, then her parents worry.

(b) Janet did not come home in time.

Question: Were her parents worried?

( ) Yes ( ) No ( ) Not enough clues

T

12. Clues: (a) If a person is not older than 16, then he does not have a driver's license.

(b) Michael does not have a driver's license.

Question: Is he older than 16?

( ) Yes ( ) No ( ) Not enough clues

T'

12. Clues: (a) If a record has no crack, then it is not John's.

(b) This record is not John's.

Question: Does it have a crack?

( ) Yes ( ) No ( ) Not enough clues

T

13. Clues: (a) If Janet does not come home in time, then her parents worry.

(b) Janet did not come home in time.

Question: Were her parents worried?

( ) Yes ( ) No ( ) Not enough clues

T'

13. Clues: (a) If Jack is not in the race, then his team will win.

(b) Jack is not in the race.

Question: Will his team win?

( ) Yes ( ) No ( ) Not enough clues

T

14. Clues: (a) If the aquarium is dirty, then the goldfish will die.

(b) The goldfish has died.

Question: Was the aquarium dirty?

( ) Yes ( ) No ( ) Not enough clues

T'

14. Clues: (a) If there is a policeman at the corner, then Don waits on the sidewalk.

(b) Don waits on the sidewalk.

Question: Is there a policeman at the corner?

( ) Yes ( ) No ( ) Not enough clues

T

15. Clues: (a) If George does not like this kind of salad, then he will try the other kind.

(b) George likes this kind of salad.

Question: Will he try the other kind?

( ) Yes ( ) No ( ) Not enough clues

T'

15. Clues: (a) If I don't see Dennis today, then I'll see him tomorrow.

(b) I've seen Dennis today.

Question: Will I see him tomorrow?

( ) Yes ( ) No ( ) Not enough clues

T

16. Clues: (a) If the coat is black, then it is not Jill's.

(b) This is Jill's coat.

Question: Is it black?

( ) Yes ( ) No ( ) Not enough clues

T'

16. Clues: (a) If it is a holiday, then the library is not open.

(b) The library is open.

Question: Is it a holiday?

( ) Yes ( ) No ( ) Not enough clues

T

17. Clues: (a) If it is Monday, then mother stays at home.

(b) It is not Monday.

Question: Is mother staying at home?

( ) Yes ( ) No ( ) Not enough clues

T'

17. Clues: (a) If father fills up his car with gas, then he cleans the windshield.

(b) Father does not fill up his car with gas.

Question: Is he cleaning the windshield?

( ) Yes ( ) No ( ) Not enough clues

T

18. Clues: (a) If a record has no crack, then it is not John's.

(b) This record is not John's.

Question: Does it have a crack?

( ) Yes ( ) No ( ) Not enough clues

T'

18. Clues: (a) If a person is not older than 16, then he does not have a driver's license.

(b) Michael does not have a driver's license.

Question: Is he older than 16?

( ) Yes ( ) No ( ) Not enough clues

T

19. Clues: (a) If it is not a children's show, then Ronald does not watch it.

(b) This is not a children's show.

Question: Does Ronald watch it?

( ) Yes ( ) No ( ) Not enough clues

T'

19. Clues: (a) If the weather is not warm, then Cindy does not go swimming.

(b) The weather is not warm.

Question: Is Cindy going to swim?

( ) Yes ( ) No ( ) Not enough clues

T

20. Clues: (a) If a peach is not soft, then it is not tasty.

(b) This peach is tasty.

Question Is it soft?

( ) Yes ( ) No ( ) Not enough clues

T'

20. Clues: (a) If the milk isn't chilled, then I'm not going to drink it.

(b) I am going to drink the milk.

Question: Is the milk chilled?

( ) Yes ( ) No ( ) Not enough clues

T

21. Clues: (a) If a student does not finish homework, then he goes to the principal's office.

(b) This student is going to the principal's office.

Question: Did he finish his homework?

( ) Yes ( ) No ( ) Not enough clues

T'

21. Clues: (a) If that woman is not Mrs. Brown, then she is Nancy's grandma.

(b) This woman is Nancy's grandma.

Question: Is she Mrs. Brown?

( ) Yes ( ) No ( ) Not enough clues

T

22. Clues: (a) If he is in room #10, then he is a 2-nd grader.

(b) He is not a 2-nd grader.

Question: Is he in room #10?

( ) Yes ( ) No ( ) Not enough clues

T'

22. Clues: (a) If it is our store, then it is on the corner.

(b) This store is not on the corner.

Question: Is it our store?

( ) Yes ( ) No ( ) Not enough clues

T

23. Clues: (a) If the wind blows from the west, then the clouds will go away.

(b) The wind blows from the west.

Question: Will the clouds go away?

( ) Yes ( ) No ( ) Not enough clues

T'

23. Clues: (a) If Mary is seven, then she is too young for that summer camp.

(b) Mary is seven.

Question: Is she too young for that summer camp?

( ) Yes ( ) No ( ) Not enough clues

T

24. Clues: (a) If their car is not in their garage, then they are not home.

(b) Their car is in their garage.

Question: Are they home?

( ) Yes ( ) No ( ) Not enough clues

T'

24. Clues: (a) If it cannot swim, then it is not a fish.

(b) It can swim.

Question: Is it a fish?

( ) Yes ( ) No ( ) Not enough clues

T

25. Clues: (a) If Dick has fruit for dessert, then he does not have cake.

(b) Dick has fruit for dessert.

Question: Does he have cake?

( ) Yes ( ) No ( ) Not enough clues

T'

25. Clues: (a) If a house has a red roof, then it is not Joy's.

(b) This house has a red roof.

Question: Is it Joy's house?

( ) Yes ( ) No ( ) Not enough clues

T

26. Clues: (a) If it cannot swim, then it is not a fish.

(b) It can swim.

Question: Is it a fish?

( ) Yes ( ) No ( ) Not enough clues

T'

26. Clues: (a) If their car is not in their garage, then they are not home.

(b) Their car is in their garage.

Question: Are they home?

( ) Yes ( ) No ( ) Not enough clues

T

27. Clues: (a) If that woman is not Mrs. Brown, then she is Nancy's grandma.

(b) This woman is Nancy's grandma.

Question: Is shw Mrs. Brown?

( ) Yes ( ) No ( ) Not enough clues

T'

27. Clues: (a) If a student does not finish homework, then he goes to the principal's office.

(b) This student is going to the principal's office.

Question: Did he finish his homework?

( ) Yes ( ) No ( ) Not enough clues

T

28. Clues: (a) If the milk isn't chilled, then I'm not going to drink it.

(b) I am going to drink the milk.

Question: Is the milk chilled?

( ) Yes ( ) No ( ) Not enough clues

T'

28. Clues: (a) If a peach is not soft, then it is not tasty.

(b) This peach is tasty.

Question: Is it soft?

( ) Yes ( ) No ( ) Not enough clues

T

29. Clues: (a) If I don't see Dennis today, then I'll see him tomorrow.

(b) I've seen Dennis today.

Question: Will I see him tomorrow?

( ) Yes ( ) No ( ) Not enough clues

T'

29. Clues: (a) If George does not like this kind of salad, then he will try the other kind.

(b) George likes this kind of salad.

Question: Will he try the other kind?

( ) Yes ( ) No ( ) Not enough clues

T

30. Clues: (a) If a house has a red roof, then it is not Joy's.

(b) This house has a red roof.

Question: Is it Joy's house?

( ) Yes ( ) No ( ) Not Enough clues

T'

30. Clues: (a) If Dick has fruit for dessert, then he does not have cake.

(b) Dick has fruit for dessert.

Question: Does he have cake?

( ) Yes ( ) No ( ) Not enough clues

T

31. Clues: (a) If it is our store, then it is on the corner.

(b) This store is not on the corner.

Question: Is it our store?

( ) Yes ( ) No ( ) Not enough clues

T'

31. Clues: (a) If he is in room #10, then he is a 2-nd grader.

(b) He is not a 2-nd grader.

Question: Is he in room #10?

( ) Yes ( ) No ( ) Not enough clues

T

32. Clues: (a) If a student takes Spanish, then he does not take French.

(b) Jeff does not take Spanish.

Question: Does he take French.

( ) Yes ( ) No ( ) Not enough clues

T'

32. Clues: (a) If the dog out there has black spots, then it is not my dog.

(b) The dog out there does not have black spots.

Question: Is it my dog?

( ) Yes ( ) No ( ) Not enough clues

APPENDIX 7.3

a. Test Versions Used in Pilot Study

b. Item profiles for version T_1

Appendix 7.3a Test Versions Used in Pilot Study
Version T_1

Full Name ______________________________

Todays Date ______________________ Team # __________

Puzzles for Fun

We are going to have a team-contest, in puzzle solving. Each student will work on his own paper. The points each member gets will sum up to the team's score.

- You will get one point for each correct answer -

- You will lose one point for each wrong answer -

Therefore, each time you try a puzzle, try hard to chose the right answer so that you and your team will earn a point. But if you are really not sure, leave the puzzle unanswered, so that your team will not lose a point. Just go to the next puzzle. If you put a sign like * next to a puzzle you skip, then you can come back to it when you finish.

If you find a word that you do not know, raise your hand, and I will explain that word to you. THIS IS NOT A TEST OF YOUR READING, so ask about any word you are not sure of.

These puzzles are just for your fun, and you may take all the time you want on them. There is NO TIME LIMIT. The competition is only on the number of right answers you get. So, work carefully and help your team win.

When you finish, check your answers again, then you may draw some pictures or design on the front page, but PLEASE, REMAIN QUIET until I collect the papers.

The Puzzles

In each puzzle we give you two clues, then we ask a question. For each question there are three possible answers: ( ) Yes, ( )No, ( )Not Enough Clues. Decide which is the right answer for each puzzle, and indicate this by marking x in the parentheses to the left of that answer.

In some puzzles we did not give you enough clues to reach a yes or a no answer. So, if you think there are not enough clues in a given puzzle, be sure to put your x next to the answer - Not enough clues. In some puzzles, of course, there are enough clues to reach a yes or a no answer. In such cases, be sure to mark your x next to yes or no.

Remember: In each case MARK ONLY ONE ANSWER

Let's do two examples together:

Example 1. Clues: (a) If it rains tomorrow, then it will be cloudy tomorrow.

(b) It will not be cloudy tomorrow.

Question: Will it rain tomorrow?

( )Yes ( )No ( )Not enough clues

The answer is: No. It will certainly not rain tomorrow, because if it does, then, according to the first clue, it will be cloudy. But the second clue says it will not be cloudy. So, No is the right answer. Mark x in the parentheses to the left of that answer.

Example 2. Clues: (a) If it rains tomorrow, then it will be cloudy tomorrow.

(b) It will not rain tomorrow.

Question: Will it be cloudy tomorrow?

( )Yes ( )No ( )Not enough clues

The right answer is: Not enough clues. The first clue says nothing about what will happen if it does not rain. It may be cloudy tomorrow even though it does not rain, or it may be not cloudy and not raining. We don't have enough clues to decide. Mark this answer.

Do you have any questions before we begin?

Now, turn the page and do your best.

1. Clues: (a) If Mary is seven, then she is too young for that summer camp.

(b) Mary is seven.

Question: Is she too young for that summer camp?

( )Yes ( )No ( )Not enough clues

2. Clues: (a) If there is a policeman at the corner, then Don waits on the sidewalk.

(b) Don waits on the sidewalk.

Question: Is there a policeman at the corner?

( )Yes ( )No ( )Not enough clues

3. Clues: (a) If the sun is not shining, then Cindy does not go swimming.

(b) The sun is not shining.

Question: Will Cindy go swimming?

( )Yes ( )No ( )Not enough clues

4. Clues: (a) If Terry has a fever, then he will not go to school tomorrow.

(b) Terry will not go to school tomorrow.

Question: Does Terry have a fever?

( )Yes ( )No ( )Not enough clues

5. Clues: (a) If Laura's desk is not straightened, then she has to stay after school.

(b) Laura does not have to stay after school.

Question: Is Laura's desk straightened?

( )Yes ( )No ( )Not enough clues

6. Clues: (a) If the dog out there has black spots, then it is not my dog.

(b) The dog out there does not have black spots.

Question: Is it my dog?

( )Yes ( )No ( )Not enough clues

7. Clues: (a) If someone plays too much football, then he does not do enough homework.

(b) Steve does not do enough homework.

Question: Does Steve play too much football?

( )Yes ( )No ( )Not enough clues

8. Clues: (a) If there is a holiday next Wednesday, then the library will not be open.

(b) The library will be open next Wednesday.

Question: Is there a holiday next Wednesday?

( )Yes ( )No ( ) Not enough clues

9. Clues: (a) If the wind does not change direction, then our sailboat will approach the dock.

(b) Our sailboat does not approach the dock.

Question: Did the wind change direction?

( )Yes ( )No ( )Not enough clues

10. Clues: (a) If father fills up his car with gas, then he cleans the windshield.

(b) Father doesn't fill up his car with gas.

Question: Is he cleaning the windshield?

( )Yes ( )No ( )Not enough clues

11. Clues: (a) If Jack is not in the race, then Joe's team will win.

(b) Jack is not in the race.

Question: Will Joe's team win?

( )Yes ( )No ( )Not enough clues

12. Clues: (a) If a person is not older than 16, then he does not have a driver's license.

(b) Michael does not have a driver's license.

Question: Is he older than 16?

( )Yes ( )No ( )Not enough clues

13. Clues: (a) If Janet does not come home in time, then her parents worry.
(b) Janet does not come home in time.
Question: Are her parents worried?
( )Yes ( )No ( )Not enough clues

14. Clues: (a) If that woman is Mrs. Brown, then she is Nancy's Grandma.
(b) That woman is Nancy's Grandma.
Question: Is she Mrs. Brown?
( )Yes ( )No ( )Not enough clues

15. Clues: (a) If George does not like this kind of salad, then he will try the other kind.
(b) George likes this kind of salad.
Question: Will he try the other kind?
( )Yes ( )No ( )Not enough clues

16. Clues: (a) If the coat is black, then it is not Jill's.
(b) This is Jill's coat.
Question: Is it black?
( )Yes ( )No ( )Not enough clues

17. Clues: (a) If it is Monday, then mother stays at home.

(b) It is not Monday.

Question: Is mother at home?

( )Yes ( )No ( )Not enough clues

18. Clues: (a) If a record has no crack in it, then it is not Jeremy's.

(b) This record is not Jeremy's.

Question: Does it have a crack in it?

( )Yes ( )No ( )Not enough clues

19. Clues: (a) If it is not a children's show, then Ronald does not watch it.

(b) This is not a children's show.

Question: Does Ronald watch it?

( )Yes ( )No ( )Not enough clues

20. Clues: (a) If a peach is not soft, then it is not tasty.

(b) This peach is tasty.

Question: Is it soft?

( )Yes ( )No ( )Not enough clues

21. Clues: (a) If a student does not finish homework, then he goes
to the principal's office.

(b) This student is going to the principal's office.

Question: Did he finish his homework?

( )Yes ( )No ( )Not enough clues

22. Clues: (a) If it is a little piece, then it will fit your puzzle.

(b) This piece does not fit your puzzle.

Question: Is it a little piece?

( )Yes ( )No ( )Not enough clues

23. Clues: (a) If the wind blows from the west, then the clouds will
go away.

(b) The wind blows from the west.

Question: Will the clouds go away?

( )Yes ( )No ( )Not enough clues

24. Clues: (a) If their car is not in their garage, then they are
not home.

(b) Their car is in their garage.

Question: Are they home?

( )Yes ( )No ( )Not enough clues

25. Clues: (a) If Dick has fruit for dessert, then he does not have cake.

(b) Dick has fruit for dessert.

Question: Does he have cake?

( )Yes ( )No ( )Not enough clues

26. Clues: (a) If this is not a mammal, then it is not a whale.

(b) This is a mammal.

Question: Is it a whale?

( )Yes ( )No ( )Not enough clues

27. Clues: (a) If the aquarium is not clean, then the goldfish will die.

(b) The goldfish has died.

Question: Was the aquarium clean?

( )Yes ( )No ( )Not enough clues

28. Clues: (a) If the milk is not chilled, then I'm not going to drink it.

(b) I am going to drink the milk.

Question: Is the milk chilled?

( )Yes ( )No ( )Not enough clues

29. Clues: (a) If I don't see him today, then I'll see him tomorrow.

(b) I've seen him today.

Question: Will I see him tomorrow.

( )Yes ( )No ( )Not enough clues

30. Clues: (a) If this house has a red roof, then it is not Joy's house.

(b) This house has a red roof.

Question: Is it Joy's house?

( )Yes ( )No ( ) Not enough clues

31. Clues: (a) If it is our store, then it is on the corner.

(b) This store is not on the corner.

Question: Is it our store?

( )Yes ( )No ( )Not enough clues

32. Clues: (a) If he takes music class, then he is not supposed to be here.

(b) This boy does not take music class.

Question: Is he supposed to be here?

( )Yes ( )No ( )Not enough clues

Version T_2

Full Name ____

Todays Date ____________________________ Team # __________

Puzzles for Fun

We are going to have a team-contest, in puzzle solving. Each student will work on his own paper. The points each member gets will sum up to the team's score.

- You will get one point for each correct answer -

- You will lose one point for each wrong answer -

Therefore, each time you try a puzzle, try hard to chose the right answer so that you and your team will earn a point. But if you are really not sure, leave the puzzle unanswered, so that your team will not lose a point. Just go to the next puzzle. If you put a sign like * next to a puzzle you skip, then you can come back to it when you finish.

If you find a word that you do not know, raise your hand, and I will explain that word to you. THIS IS NOT A TEST OF YOUR READING, so ask about any word you are not sure of.

These puzzles are just for your fun, and you may take all the time you want on them. There is NO TIME LIMIT. The competition is only on the number of right answers you get. So, work carefully and help your team win.

When you finish, check your answers again, then you may draw some pictures or design on the front page, but PLEASE, REMAIN QUIET until I collect the papers.

The Puzzles

In each puzzle we give you two clues, then we ask a question. For each question there are three possible answers: ( ) Yes, ( )No, ( )Not Enough Clues. Decide which is the right answer for each puzzle, and indicate this by marking x in the parentheses to the left of that answer.

In some puzzles we did not give you enough clues to reach a yes or a no answer. So, if you think there are not enough clues in a given puzzle, be sure to put your x next to the answer - Not enough clues. In some puzzles, of course, there are enough clues to reach a yes or a no answer. In such cases, be sure to mark your x next to yes or no.

Remember: In each case MARK ONLY ONE ANSWER

Let's do two examples together:

Example 1. Clues: (a) If it rains tomorrow, then it will be cloudy tomorrow.

(b) It will not be cloudy tomorrow.

Question: Will it rain tomorrow?

( )Yes ( )No ( )Not enough clues

The answer is: No. It will certainly not rain tomorrow, because if it does, then, according to the first clue, it will be cloudy. But the second clue says it will not be cloudy. So, No is the right answer. Mark x in the parentheses to the left of that answer.

Example 2. Clues: (a) If it rains tomorrow, then it will be cloudy tomorrow.

(b) It will not rain tomorrow.

Question: Will it be cloudy tomorrow?

( )Yes ( )No ( )Not enough clues

The right answer is: Not enough clues. The first clue says nothing about what will happen if it does not rain. It may be cloudy tomorrow even though it does not rain, or it may be not cloudy and not raining. We don't have enough clues to decide. Mark this answer.

Do you have any questions before we begin?

Now, turn the page and do your best.

1. Clues: (a) If Mary is seven, then she is too young for that summer camp.

(b) Mary is not too young for that summer camp.

Question: Is Mary seven?

( )Yes ( )No ( )Not enough clues

2. Clues: (a) If there is a policeman at the corner, then Don waits on the sidewalk.

(b) There is no policeman at the corner.

Question: Will Don wait on the sidewalk?

( )Yes ( )No ( )Not enough clues

3. Clues: (a) If the sun is not shining, then Cindy does not go swimming.

(b) Cindy goes swimming.

Question: Is the sun shining?

( )Yes ( )No ( )Not enough clues

4. Clues: (a) If Terry has a fever, then he will not go to school tomorrow.

(b) Terry does not have a fever.

Question: Will he go to school tomorrow?

( )Yes ( )No ( )Not enough clues

5. Clues: (a) If Laura's desk is not straightened, then she'll have to stay after school.

(b) Laura's desk is not straightened.

Question: Will she have to stay after school?

( )Yes ( )No ( )Not enough clues

6. Clues: (a) If the dog out there has black spots, then it is not my dog.

(b) This is not my dog.

Question: Does it have black spots?

( )Yes ( )No ( )Not enough clues

7. Clues: (a) If someone plays too much football, then he does not do enough homework.

(b) Steven does not play too much football.

Question: Does he do enough homework?

( )Yes ( )No ( )Not enough clues

8. Clues: (a) If there is a holiday next Wednesday, then the library will not be open.

(b) There is a holiday next Wednesday.

Question: Will the library be open?

( )Yes ( )No ( )Not enough clues

9. Clues: (a) If the wind does not change direction, then our sailboat will approach the bank.

(b) The wind did not change direction.

Question: Does our sailboat approach the bank?

( )Yes ( )No ( )Not enough clues

10. Clues: (a) If father fills up his car with gas, then he cleans the windshield.

(b) Father cleaned the windshield.

Question: Did he fill up his car with gas?

( )Yes ( )No ( )Not enough clues

11. Clues: (a) If Jack is not in the race, then Joe's team will win.

(b) Joe's team did not win.

Question: Was Jack in the race?

( )Yes ( )No ( )Not enough clues

12. Clues: (a) If a person is not older than 16, then he does not have a driver's license.

(b) Michael is older than 16.

Question: Does he have a driver's license?

( )Yes ( )No ( )Not enough clues

13. Clues: (a) If Janet does not come home in time, then her parents worry.

(b) Janet's parents are not worried.

Question: Did she come home in time?

( )Yes ( )No ( )Not enough clues

14. Clues: (a) If that woman is Mrs. Brown, then she is Nancy's grandma.

(b) That woman is not Mrs. Brown.

Question: Is she Nancy's grandma?

( )Yes ( )No ( )Not enough clues

15. Clues: (a) If George does not like this kind of salad, then he will try the other kind.

(b) George is trying the other kind of salad.

Question: Does he like the first kind of salad?

( )Yes ( )No ( )Not enough clues

16. Clues: (a) If the coat is black, then it is not Jill's.

(b) This coat is black,

Question: Is it Jill's?

( )Yes ( )No ( )Not enough clues

17. Clues: (a) If it is Monday, then mother stays at home.

(b) Mother stays home.

Question: Is it Monday?

( )Yes ( )No ( )Not enough clues

18. Clues: (a) If a record has no crack in it, then it is not Jeremy's.

(b) This record has a crack in it.

Question: Is it Jeremy's.

( )Yes ( )No ( )Not enough clues

19. Clues: (a) If it is not a children's show, then Ronald does not watch it.

(b) Ronald watches a show.

Question: Is it a children's show?

( )Yes ( )No ( )Not enough clues

20. Clues: (a) If a peach is not soft, then it is not tasty.

(b) This peach is not soft.

Question: Is it tasty?

( )Yes ( )No ( )Not enough clues

21. Clues: (a) If a student does not finish homework, then he goes to the principal's office.

(b) This student finished his homework.

Question: Does he go to the principal's office?

( )Yes ( )No ( )Not enough clues

22. Clues: (a) If it is a little piece, then it will fit your puzzle.

(b) This is a little piece.

Question: Will it fit your puzzle?

( )Yes ( )No ( )Not enough clues

23. Clues: (a) If the wind blows from the west, then the clouds will go away.

(b) The clouds do not go away.

Question: Does the wind blow from the west?

( )Yes ( )No ( )Not enough clues

24. Clues: (a) If their car is not in their garage, then they are not home.

(b) They are not home.

Question: Is their car in their garage?

( )Yes ( )No ( )Not enough clues

25. Clues: (a) If Dick has fruit for dessert, then he does not have cake.

(b) Dick has cake for dessert.

Question: Does he have fruit?

( )Yes ( )No ( )Not enough clues

26. Clues: (a) If this is not a mammal, then it is not a whale.

(b) This is not a whale.

Question: Is it a mammal?

( )Yes ( )No ( )Not enough clues

27. Clues: (a) If the aquarium is not clean, then the goldfish will die.

(b) The aquarium is clean.

Question: Will the goldfish die?

( )Yes ( )No ( )Not enough clues

28. Clues: (a) If the milk is not chilled, then I'm not going to drink it.

(b) The milk is not chilled.

Question: Am I going to drink it?

( )Yes ( )No ( )Not enough clues

29. Clues: (a) If I don't see him today, then I'll see him tomorrow.

(b) I'll see him tomorrow.

Question: Have I seen him today?

( )Yes ( )No ( )Not enough clues

30. Clues: (a) If this house has a red roof, then it is not Joy's house.

(b) Here is Joy's house.

Question: Does it have a red roof?

( )Yes ( )No ( )Not enough clues

31. Clues: (a) If it is our store, then it is on the corner.

(b) This is our store.

Question: Is it on the corner?

( )Yes ( )No ( )Not enough clues

32. Clues: (a) If he takes music class, then he is not supposed to be here.

(b) This boy is not supposed to be here.

Question: Does he take music class?

( )Yes ( )No ( )Not enough clues

Appendix 7.3b Item Profiles in Percentage of Students Per Answer for Pilot Study Students Who Took Version T_1 Either as Their Pretest or as Their Posttest

Item No.	Logical Form	Negation Mode	T_1 taken in pretest						T_1 taken in posttest					
			Experimental (n=24)			Control (n=38)			Experimental (n=25)			Control (n=37)		
			Yes	No	NEC	Yes	No	NEC	Yes	No	NEC	Yes	No	NEC
1	MP	++	75.9	16.7	8.3	78.9	5.3	15.8	84.0	8.0	8.0	86.5	10.8	2.7
23		++	91.7	8.3	.0	81.6	5.3	13.2	76.0	.0	24.0	89.2	.0	10.8
25		+-	.0	100.0	.0	2.6	97.4	.0	.0	72.0	28.0	2.7	89.2	8.1
30		+-	37.5	41.7	20.8	36.8	52.6	10.5	36.0	36.0	28.0	32.4	54.1	13.5
11		-+	95.8	4.2	.0	81.6	7.9	10.5	68.0	8.0	24.0	64.9	10.8	24.3
13		-+	100.0	.0	.0	89.5	2.6	7.9	68.0	.0	32.0	83.8	8.1	8.1
19		--	16.7	79.2	4.2	21.1	78.9	.0	8.0	60.0	32.0	13.5	70.3	16.2
3		--	16.7	83.3	.0	10.5	84.2	5.3	8.0	68.0	24.0	5.4	81.1	13.5
22	MT	++	8.3	70.8	20.8	7.9	84.2	7.9	.0	64.0	36.0	21.6	54.1	24.3
31		++	16.7	83.3	.0	5.3	86.8	7.9	8.0	80.0	12.0	8.1	83.8	8.1
8		+-	16.7	70.8	12.5	21.1	60.5	18.4	8.0	76.0	16.0	.0	86.5	13.5
16		+-	25.0	62.5	12.5	31.6	68.4	.0	24.0	56.0	20.0	21.6	64.9	13.5
5		-+	62.5	25.0	12.5	71.1	21.1	7.9	76.0	4.0	20.0	89.2	5.4	5.4
9		-+	54.2	29.2	16.7	52.6	39.5	7.9	32.0	16.0	52.0	56.8	27.0	16.2
20		--	58.3	29.2	12.5	76.3	23.7	.0	64.0	12.0	24.0	64.9	29.7	5.4
28		--	75.0	25.0	.0	68.4	23.7	7.9	84.0	4.0	12.0	78.4	16.2	5.4
2	AC	++	50.0	20.8	29.2	52.6	23.7	23.7	20.0	.0	80.0	67.6	5.4	27.0
14		++	79.2	12.5	8.3	78.9	10.5	10.5	56.0	4.0	40.0	83.8	2.7	13.5
4		+-	70.8	8.3	20.8	68.4	10.5	21.1	28.0	4.0	68.0	62.2	8.1	29.7
7		+-	75.0	.0	25.0	89.5	2.6	7.9	28.0	.0	72.0	67.6	2.7	29.7
21		-+	8.3	70.8	20.8	2.6	89.5	7.9	4.0	28.0	68.0	5.4	62.2	32.4
27		-+	.0	79.2	20.8	2.6	86.8	10.5	4.0	52.0	44.0	2.7	75.7	21.6
12		--	8.3	83.3	8.3	7.9	78.9	13.2	.0	48.0	52.0	13.5	54.1	32.4
18		--	25.0	58.3	16.7	31.6	55.3	13.2	4.0	24.0	72.0	27.0	43.2	29.7
17	DA	++	8.3	66.7	25.0	10.5	81.6	7.9	16.0	32.0	52.0	5.4	73.0	21.6
10		++	33.3	37.5	29.2	26.3	5.0	23.7	4.0	36.0	60.0	10.8	48.6	40.5
6		+-	33.3	50.0	16.7	47.4	34.2	18.4	16.0	20.0	64.0	32.4	32.4	35.1
32		+-	45.8	37.5	16.7	52.6	34.2	13.2	8.0	52.0	40.0	37.8	35.1	27.0
29		-+	8.3	58.3	33.3	5.3	65.8	28.9	4.0	20.0	76.0	2.7	35.1	62.2
15		-+	12.5	70.8	16.7	18.4	63.2	18.4	4.0	28.0	68.0	2.7	64.9	32.4
24		--	62.5	20.8	16.7	78.9	5.3	15.8	32.0	12.0	56.0	70.3	2.7	27.0
26		--	54.2	20.8	25.0	44.7	39.5	15.8	24.0	28.0	48.0	35.1	29.7	35.1

APPENDIX 7.4

a. Students' Attitude Questionnaire

b. Teachers' Evaluation of the Experimental Unit Questionnaire

c. Teachers' Evaluation of Activities Questionnaire

Appendix 7.4a Students' Attitude Questionnaire

(taken by experimental group students right after the posttest)

1. In the last 5 weeks you worked on the first logic project. We would like to know which parts of it you like, and which parts you did not. Here is a list of the activities to remind you what you did. As you read it circle the ones you really liked:

   a. Electric cards

   b. Pictorial activity

   c. Dominoes

   d. Numbers and their properties

   e. Playing cards

   f. Colored light switch box

   g. Prepare a quiz

   Now, please go again and cross the ones you did not like at all.

2. Right after Easter vacation we may offer a new logic project. Only Those students who choose to participate in it will take it. It's up to you. Would you like to go on learning logic after Easter?________

   Do you have a reason for it?________________________________

   ____

   ____

Appendix 7.4b

Teacher's overall evaluation of experimental unit in conditional reasoning

March 1975

1. Teacher's name (optional) ______________________________

2. No. of years of elementary school teaching experience __________

How many years in present school? ________

3. A.B. degree: College ______________________________

Year ________________

Major ________________

Credential program: University: ______________________

Year ____________

Other graduate work: ______________________________

____

4. Did you take any logic courses? _______. If yes, state where, and give a brief description of the course: ____________________

____

5. Have you ever been involved in educational research? _________

If yes, in what way? ______________________________

____

____

____

6. Name professional journals you read regularly __________

____

____

____

7. Here is a list of the activities the experimental unit consists of:

Electric cards, Pictorial activity, Dominoes, Numbers and their properties, Playing cards, Colored light switch box, Prepare a quiz.

Which activity did you like best? ____________________

Why? ____________________

Which activity did you like the least ____________________

Why? ____________________

8. Which activity do you think your class liked the most? ____________________

Least? ____________________

9. Which activity was the most helpful towards understanding of the underlying logical ideas? ____________________

Which activities caused confusion in the students' minds about the underlying logical ideas? ____________________

10. Please comment on the sequencing of the activities. (Any changes you think would be better) ____________________

11. Would you like to teach this unit again next year? ____________

Please give reason to either a yes or a no answer ____________

12. If your answer to item 11 was yes, go to 13. If your answer to item 11 was no, do you think some additional training might change your mind? ____________

13. When you teach this unit again, would you modify the approach? ______

How? ____________

14. About how much time did you spend in preparation for a class?

Did you read the teacher's manual in full? ________ How did it serve you?(Is it too long? too detailed? Comment on it's editting, language, clearity, etc.) ____________

15. Do you think you would have been able to teach this unit using just the teacher's manual, with no pre-training session? ____________

If not, do you think it would be possible to improve the manual so that it would be the sole source of instruction for the teacher? ___

(Any idea or suggestion about such an improvement will be welcomed)

____

____

____

____

16. When you prepared for a session, for which activity of this unit did you feel that your training was not adequate. (Please refer to the list in question 7 . Also, please explain what was your trouble)

____

____

____

____

17. Please comment on the pre-training session (6 meetings during January) ____

____

____

____

____

____

18. Do you think the rate of progress was approximately uniform throughout the 5-weeks program? ________ If not, which were the periods of greater learning? ______________________________

Of least learning? ______________________________

Give any explanation that occurs to you to explain such differences

____

____

____

____

19. Please comment on changes in students' attitude thruoghout the 5-weeks period. Give any explanation that occurs to you for it.

____

____

____

____

____

20. Can you identify types of students for whom this unit was particularly appropriate? ______________________________

____

Inappropriate? ______________________________________

____

21. To what extent do ability groups with respect to the work in this unit, conform to ability groups with respect to mathematics?

____

Language arts? ______________________________________

____

General ability? ______________________________

Give any explanation that occurs to you ______________________________

22. Do you think this unit may have any carry over effect on the students' work in other parts of the curriculum? If yes, what parts? ________

23. Do you think this unit may affect your teaching of regular math program? ________ how? ______________________________

Language arts? ________ How? ______________________________

Other? How? ______________________________

24. Any other coments? ______________________________

Thank you very much for your cooperation

Appendix 7.4c

Teacher's evaluation of activities

Teacher's name (Optional) ____________________

Date: ____________________

Activity evaluated:

___ Electric cards
___ Numbers and their properties
___ Pictorial activity
___ Numbers and their propertoes
___ Playing cards
___ Colored light switch box
___ Prepare a quiz

1. Which parts of that activity did you implement in your class? (State pages numbers in teacher's manual, please.) ____________________

2. What changes did you make in the parts you implemented, and why did you do them? (Youmay use the back of this paper if you need extra space. please be very explicit) ____________________

3. Which parts of that activity did you omit, and why? (State pages numbers in teacher's manual, use the back of this page for extra space)

4. Did you think of things to do that were not mentioned at all in the teacher's manual? If so, indicate what these are and whether you actually tried them. If you did, how did they work?

5. Comment on students' response and invovement in this activity.(Quote interesting reactions, use an extra sheet of paper, if needed)

6. Comment on students' worksheets and/or manipulative aids of this activity. (State items that were too easy, too difficult, boring, exciting, complicated etc., What did you do to overcome such problems?)

7. Comment on teacher's manual and answers sheets for this activity. (Did you have any parts that you felt you were not well prepared for their presentation? What parts in the manual you did not read, or you felt were not needed? etc.)

APPENDIX 7.5

A Session by Session Account of the Pretraining Workshop

January 14, 1975:[*]
1. General goal of the experimental unit.
2. Electric Cards - 20 cards: read, answer, discuss and test with the electric tester.
3. Pictorial Activity part a. (Last part assigned for homework.)

Handouts: Electric Cards - teacher's manual.
Pictorial Activity - students' worksheets.

January 16, 1975:
1. The structure of a conditional sentence and its symbolic representation.
2. Pictorial Activity part b.

d. Dominoes Activities.

4. Generalization to: $p \rightarrow q$ implies $-(p \wedge -q)$

Handouts: Pictorial Activity and Dominoes - teacher's manual.

January 21, 1975:
1. Numbers and Their Properties, including students' worksheets.
2. Playing Cards (students' worksheets assigned for homework.)
3. The meaning of NEC as a denial of both a certain yes and a certain no.
4. Generalization: $p \rightarrow q$ is equivalent to $-(p \wedge -q)$; $p \rightarrow q$ is equivalent to $(p \wedge q) \vee (-p \wedge q) \vee (-p \wedge -q)$

Handouts: Numbers and Their Properties - teachers' manual. Playing Cards - students' worksheets.

---

[*] This list includes only the new activities and ideas discussed each time. It should be understood that at many points there were reviews of accumulated knowledge.

January 23, 1975:
1. Playing Cards.
2. Colored Light Switch Box.
3. Generalization: $p \to q$ implies not-q → not-p. $p \to q$ and $q \to p$ are independent.

Handouts: Playing Cards, Colored Light Switch Box, The Logic of Conditional Reasoning - teacher's manual.

January 28, 1975:
1. Prepare a Quiz.
2. The Logic of Conditional Reasoning - summary of logical forms and negative modes through the way the electric cards were organized.
3. Psychological studies of the difficulties in conditional reasoning.

February 4, 1975:
1. Practical problems of presenting the unit in class - grouping, insisting on sound arguments for answers, encouraging students' discoveries of patterns without any teaching of algorithmic approach, equal emphasis on all four logical forms, gradual introduction of negation into conditional sentences.
2. Teachers' pretest, self corrected.

BIBLIOGRAPHY

The following list includes books and papers consulted as well as ones explicitly referred to. The latter are marked by an asterisk.

Allen, L. E. (1965). "Toward autotelic learning of mathematical logic by the WFF'N PROOF game." In L. N. Morrisett and J. Vinsohaler (Eds.) *Mathematical Learning*. Monographs of the Society for Research in Child Development, 30:1, 29-41.

Almy, M. (1970). *Logical Thinking in Second Grade*. Teachers College Press, New York.

*Anastasi, A. (1968). *Psychological Testing*. Third Edition. New York: Macmillan Company.

*Begg, I., and Denny, J. P. (1969). "Empirical reconciliation of atmosphere and conversion interpretations of syllogistic reasoning errors." *Journal of Experimental Psychology*, 81, 351-354.

*Brownell, W. A. (1936). "Theoretical aspects of learning and transfer of training." *Review of Educational Research*, VI.

Bruner, J. S., Goodnow, J. J., and Austin, G. A. (1956). *A Study of Thinking*. New York: Wiley.

*Burt, C. (1919). "The development of reasoning in school children." *The Journal of Experimental Pedagogy*, 68-77, 121-127.

*Campbell, D. T., and Stanley, J. C. (1963). *Experimental and Quasi-Experimental Designs for Research*. Rand McNally Company.

*Carroll, C. A. (1970). "Low Achievers' Understanding of Four Logical Inference Forms: An Analysis of Difficulties and of the Effect of Instruction." Unpublished Ph.D. dissertation. New York: Columbia University.

*CEMREL (1972). see Exner, Karmos, Kaufman, and Steiner.

*Chapman, L. J., and Chapman, J. P. (1959). "Atmosphere effect re-examined." *The Journal of Experimental Psychology*, 58, 220-226.

*Chen, B. (1975). "A Study of the Development of the Ability to Make Legitimate Inferences" (tentative title). Ph.D. dissertation in preparation. Berkeley: University of California.

*Church, A. (1956). *Introduction to Mathematical Logic*. New Jersey: Princeton University Press.

*Cohen, M. R., and Nagel, E. (1957). An Introduction to Logic and Scientific Method. London: Routledge and K. Paul.

Covington, M. V., Crutchfield, R. S., Davies, L., Olton, R. M. (1972). The Productive Thinking Program, A Course in Learn[illegible] to Think. Columbus Ohio: Ch. E. Merrill.

Cronbach, L. J., and Furby, L. (1970). "How should we measure 'change' or should we?..." Psychological Bulletin, vol. 74, No. 1, 68-80.

*CTBS-Comprehensive Test of Basic Skills (1968-1970). New York: CTB - McGraw Hill Book Company.

*Davis, R. (1964). Discovery in Mathematics. California: Addison-Wesley.

Dinnes, Z. P., and Golding, E. W. (1970). Learning Logic, Logical Games. New York: Herder and Herder.

*Donaldson, M. A. (1963). Study of Children's Thinking. London: Tavistok Publication.

*Eisenberg, T. A., and McGinty R. L. (1974). "On comparing error patterns and the effect of maturation in a unit on sentential logic." Journal of Research in Mathematics Education, Vol. 5, No. 4, 225-237.

*Elementary Science Study (ESS), (1968). Attribute Games and Problems. New York : McGraw Hill Book Company.

*Ellis, H. (1965). The Transfer of Learning. New York: Macmillan Company.

*Enderton, H. G. (1972). A Mathematical Introduction to Logic. New York: Academic Press.

Ennis, R. H. (1975). "Children's ability to handle Piaget's propositional logic: A conceptual critique." Review of Educational Research, Vol. 45, No. 1, 1-41.

*Ennis, R. H., and Paulus, D. H. (1965). Critical Thinking Readiness in Grades 1-12, Phase I: Deductive Logic in Adolescence. Ithaca: Cornell University.

Evans, J. St. Bt. B. T. (1972). "Deductive Reasoning and Linguistic Usage (with special reference to negation)." Unpublished Ph.D. dissertation, University of London.

Exner, R. M., and Rosskopf, M. F. (1959). Logic in Elementary Mathematics. New York: McGraw Hill Book Company.

*Exner, R. M., Karmos, J. S., Kaufman, B. A., and Steiner, H. G. (1972). Elements of Mathematics. Book 1: Introductory Logic. Comprehensive School Mathematics Program (CSMP). St. Ann, Missouri: CEMREL, Inc.

*Feather, N. T. (1964). "Acceptance and rejection of arguments in relation to attitude strength, critical ability and intolerance of inconsistency." Journal of Abnormal Social Psychology LXIX, 127-137.

Fey, J. (1968). "Patterns of Verbal Communication in Mathematics Classes." Unpublished Ph.D. dissertation. New York: Teachers College, Columbia University.

*Frase, L. T. (1966). "Belief, incongruity, and syllogistic reasoning." Psychological Reports, 18, 982.

Frase L. T. (1968). "Associative factors in syllogistic reasoning." Journal of Experimental Psychology, 76, 407-412.

Freudenthal, H. (1970). Mathematics as an Educational Task. Holland: D. Reidel Publishing Company.

Furth, H. C. (1970). Piaget for Teachers. Englewood Cliffs, New Jersey: Prentice-Hall.

Gagné, R. M. (1963). "Learning and proficiency in mathematics." Mathematics Teacher 56, 620-626.

*Gardiner, W. (1965). "An Investigation of Understanding the Meaning of the Logical Operators in Propositional Reasoning." Unpublished Doctoral dissertation, Cornell University.

*Gregory, J. W. (1972). "A study of the impact of the verbal environment in mathematics classroom on seventh grade students' logical abilities." Columbus, Ohio: Ohio State University. (ERIC ED 064178).

*Gregory, J. W., and Osborne, A. R. (1975). "Logical reasoning ability and teacher verbal behavior within the mathematics classroom." Journal of Research in Mathematics Education Vol. 6, No. 1, 26-36.

*Guilford, J. P., and Fruchter, B. (1973). Fundamental Statistics in Psychology and Education. San Francisco: McGraw-Hill.

Henle, M. (1968). "On the relation between logic and thinking." In P. C. Wason and P. N. Johnson-Laird (Eds.) Thinking and Reasoning, Penguin.

*Henkin, L. (1974). Private communication. Berkeley: University of California. Department of Mathematics.

*Hill, S. A. (1961). "A Study of the Logical Abilities of Children." Unpublished Ph.D. dissertation. California: Stanford University.

*Hyram, G. H. (1957). "An experiment in developing critical thinking in children." *Journal of Experimental Education* XXVI, 125-132.

Johnson, D. M. (1966). "Reasoning and logic." *Encyclopedia of Social Sciences.*

Johnson-Laird, P. N., and Tagart, J. (1969). "How implication is understood." *American Journal of Psychology*, 82, 367-373.

*Kane, R. B. (1975). "On the proof-making task." *The Mathematics Teacher*, Vol. 68, No. 2, 89-94.

*Krygowska, A. Z. (1971). "Treatment of axiomatic methods in class." In Servais, W., and Varga, T. (Eds.) *Teaching School Mathematics*. Penguin Book, Unesco.

*Lefford, A. (1946). "The influence of emotional subject matter on logical reasoning." *The Journal of General Psychology*, 34, 127-151.

*Legrenzi, P. (1971). "Relations between language and reasoning about deductive rules." In Flors D'Arcais, G. B., and Levett, W. J. M. (Eds.) *Advances in Psycholinguistics*, Amsterdam: North-Holland.

*Lester, F. K. (1975). "Developmental aspects of children's ability to understand mathematical proof." *Journal of Research in Mathematics Education*, Vol. 6, No. 1, 14-25.

Lowery, L. (1974). *Learning about Learning: Propositional Abilities*. Berkeley: University of California.

Lunzer, E. A. (1972). "The development of formal reasoning: some recent experiments and their implications." In Karl Freg und Manfred Lang Herausgeber (Eds.) *Kognitonspsychologie und naturwissenschaftliches Unterricht*. Stuttgart: Verlag Hans Huber.

*MacGinitie, W. H., and Ball, S. (1968). *Readings in Psychological Foundations of Education*. New York: McGraw-Hill Book Company.

*Marascuilo, L. A., and McSweeney, M. (in preparation). *Nonparametric and Distribution-Free Methods for Social Sciences.* Brooks-Cole, Monterey, California.

*McAloon, S. M. (1969). "An Exploratory Study on Teaching Units in Logic to Grades Three to Six." Unpublished doctoral dissertation. Ann Arbor: University of Michigan.

*McCoy, R. E. (1971). "A Study of the Effects of Three Different Strategies of Proof-Instruction and Background Factors of Elementary Education Majors, for Success in Constructing Deductive Proofs in Mathematics."

*Mendelson, E. (1964). Introduction to Mathematical Logic. Van Nostrand Company.

*Miller, W. A. (1968). "The Acceptance and Recognition of Six Logical Inference Patterns by Secondary Students." Unpublished Ph.D. dissertation, University of Wisconsin.

Miller, W. A. (1969). "Unit in sentential logic for junior high school students involving both valid and invalid inference patterns." School Science and Math, 69, 538-542.

*Morgan, J. J. B., and Carrington, D. H., (1944). "Graphic instruction in relational reasoning." Journal of Educational Psychology XXXV, 536-544.

Morgan, W. H. (1970). "A Study of the Ability of College Mathematics Students in Proof Related Logic." Unpublished Ph.D. dissertation, University of Georgia.

Nedelsky, L. (1965). Science Teaching and Testing. San Francisco: Harcourt, Brace & World, Inc.

*O'Brien, T. C. (1970). "Logical thinking in children ages 6-13." Child Development, 41, 823-829.

*O'Brien, T. C. (1973). "Logical thinking in college students." Educational Studies in Mathematics, 5, 71-79.

O'Brien, T. C. (1975). "Deformation and the four cards problem." Educational Studies in Mathematics, 6, 23-29.

*O'Brien, T. C., and Reali, N. C. (1971). "Logical thinking -- language and context." Educational Studies in Mathematics, 4, 201-219.

*O'Brien, T. C., and Shapiro, B. J. (1968). "The development of logical thinking in children." American Educational Research Journal, 531-542.

*Osherson, D. N., (1974). Logical Abilities in Children. Vol. 2, Logical Inference: Underlying Operations. Potomac, Maryland: Lawrence Erlbaum Associates, Publishers.

*Parrott, G. L. (1967). "The Effect of Premise Content on Accuracy and Solution Time in Syllogistic Reasoning." Unpublished Master's thesis. Michigan State University.

*Paulus, D. H. (1967). "A Study of Children's Ability to Deduce and to Judge Deductions." Unpublished Ph.D. dissertation, Cornell University.

*Payne, D. A. (1968). The Specification and Measurement of Learning Outcomes. U.S.A.: R. Blaisdell Publishing Company.

*Pezzoli, J. A., and Fraze, L. T. (1968). "Mediated facilitation of syllogistic reasoning." Journal of Experimental Psychology, 78, 228-232.

Piaget, J. (1957). Logic and Psychology. New York: Basic Books.

*Piaget, J. (1959). Judgment and Reasoning in the Child. Paterson, New Jersey: Littlefield, Adams & Co.

Piaget, J. (1968). Judgment and Reasoning in the Child. Totowa, New Jersey: Littlefield, Adams & Co.

*Piaget, J., and Inhelder, B. (1958). The Growth of Logical Reasoning from Childhood to Adolescence. New York: Basic Books.

Polya, G. (1954). Mathematics and Plausible Inference. Vol. 2: Patterns of Plausible Inference. New Jersey: Princeton University Press.

*Quine, W. O. (1951). Mathematical Logic. Boston: Harvard University Press.

*Roberge, J. J. (1969). "Negation in the major premise as a factor in children's deductive reasoning." School Science and Mathematics 69(8), 715-722.

*Roberge, J. J. (1970). "A study of children's abilities to reason with basic principles of deductive reasoning." American Educational Research Journal, 583-596.

Roy, G. R. (1970). "The Effect of the Study of Mathematical Logic on Students Performance in Proving Theorems by Mathematical Induction." Unpublished doctoral dissertation. Buffalo: State University of New York.

Scandura, J. M. (1967). "Learning verbal and symbolic statements of mathematical rules." Journal of Educational Psychology. 58, 356-64.

*Sells, S. B. (1936). "The atmosphere effect: An experimental study of reasoning." Archives of Psychology, 29, 3-72.

*Siegel, S. (1956). Nonparametric Statistics for the Behavioral Sciences. New York: McGraw-Hill Book Company.

*Simpson, M. E., and Johnson, D. M. (1966). "Atmosphere and conversion errors in syllogistic reasoning." Journal of Experimental Psychology, 72, 197-200.

*Siwek, H. (1973). "Logique formelle et raisonnement natural, des elèves dans l'enseignement de la mathématique." *Educational Studies in Mathematics*, 5, 23-27.

Smedsland, J. (1964). "Concrete reasoning: A study of intellectual development." *Monographs of the Society for Research in Child Development*, 29, No. 2, Ser. 93.

*Stabler, E. R. (1953). *An Introduction to Mathematical Thought*. Massachusetts: Addison-Wesley Publishing Company, Inc.

*Stanford Achievement Test (1973 edition). San Francisco: Harcourt, Brace & World, Inc.

*Stephens, J. M. (1963). "Transfer of learning." In R. Grove and R. Burney (Eds.) *Transfer of Learning*. New Jersey: D. Van-Nostrand Company, Inc.

*Suppes, P. (1957). *Introduction to Logic*. New Jersey: D. Van-Nostrand Company.

*Suppes, P. (1962). "Mathematical logic for schools." *The Arithmetic Teacher*, IX, 396-399.

*Suppes, P. (1965). "Towards a behavioral foundation of mathematical proofs." *Technical Report No. 44, Psychology Series*. Institute for Mathematical Studies in the Social Sciences, Stanford University.

*Suppes, P. and Binford, F. (1965). "Experimental teaching of mathematical logic in the elementary school." *The Arithmetic Teacher*, 12(3), 823-829.

*Suppes, P., and Feldman, S. (1971). "Young children's comprehension of logical connectives." *Journal of Experimental Child Psychology*, 12, 304-317.

*Suppes, P., and Hill, S. (1964). *First Course in Mathematical Logic*. New York: Blaisdell Publishing Company.

Swain, R. L. (1963). "Logic: For teacher, for pupil." In *Enrichment Mathematics for the Grades*. 27th Yearbook of the National Council of Teachers of Mathematics. Washington D.C.: The council.

Tabach, S. (1971). "The development of the concept of mathematical proof." In Rosskopf, M. F., and Steffe, L. P. (Eds.) *Piagetian Cognitive Development Research and Mathematical Education*. Washington D.C.: The National Council of Teachers of Mathematics.

Thorndike, E. L., and Lorge, F. (1944). *The Teachers Word Book of 30,000 Words*. New York: Teachers College Columbia University.

Tyler, R. W., Gagné, R. M. and Scriven, M. (1967). Perspectives of Curriculum Evaluation. AERA Monograph Series on Curriculum Evaluation No. 1. U.S.A.: Rand McNally.

*Vigotsky, L. S. (1965). Thought and Language. Boston: M.I.T. Press.

Wason, P. C. (1959). "The processing of positive and negative information." Quarterly Journal of Experimental Psychology, 11, 92-107.

*Wason, P. C. (1964). "The effect of self-contradiction on fallacious reasoning." Quarterly Journal of Experimental Psychology, 16, 30-34.

Wason, P. C. (1968). "Reasoning about a rule." Quarterly Journal of Experimental Psychology, 20, 273-281.

Wason, P. C., and Johnson-Laird, P. N. (1970). "A conflict between selecting and evaluating information in an inferential task." British Journal of Psychology, 61, 4, 509-515.

*Wason, P. C., and Johnson-Laird, P. N. (1972). Psychology of Reasoning, Structure and Content. Cambridge, Massachusetts: Harvard University Press.

*Weeks, G. (1970). "The Effect of Attribute Blocks Training on Second and Third Graders' Logical and Perceptual Reasoning Ability." Unpublished Ph.D. dissertation. University of Georgia.

Weiss, C. H. (1972). Evaluation Research: Methods of Assessing Program Effectiveness. New Jersey: Prentice-Hall.

Wheeier, D. (1958). "Studies in the development of reasoning in school children." British Journal of Statistical Psychology, Vol. XI, Part II.

*White, E. E. (1936). "A study of the possibility of improving habits of thought in school children by training in logic." British Journal of Educational Psychology, VI, 267-273.

Wilkins, M. C. (1928). "The effect of changed material on the ability to do formal syllogistic reasoning." Archive of Psychology, 16, 83.

Williams, B. R. (1971). "Critical Thinking Ability as Affected by a Unit on Symbolic Logic." Unpublished doctoral dissertation. Arizona State University.

Winch, W. H. (1921). "Children's reasoning: Experimental studies of reasoning in school children." Journal of Experimental Psychology, 6, 121-141, 199-212, 275-287.

Winch, W. H. (1972). Mathematics Framework for California Public Schools K-8. Strand 8: Logical Thinking. The second strands report.

*Woodworth, R. S., and Sells, S. B. (1935). "An atmosphere effect in formal syllogistic reasoning." Journal of Experimental Psychology, 18, 451-460.